langues indo-européennes

Sous la direction de
Françoise BADER





LANGUES INDO-EUROPÉENNES

COLLECTION SCIENCES DU LANGAGE

Ouvrages parus :

Emilio BONVINI : *Prédication et Enonciation en kasim.*
Danièle GODARD : *La Syntaxe des relatives en français.*
Boyd MICHAILOVSKY : *La Langue Hayu.*
Richard ZUBER : *Implications sémantiques dans les langues naturelles.*
Jean-Paul CHAUVEAU : *Evolutions phonétiques en gallo.*
Francis JOUANNET (dir.) : *Modèles en tonologie* (Kirundi et Kinyarwanda).
Jean LALLOT : *La Grammaire de Denys le Thrace* (Traduction annotée).
Béatrice GODARD-WENDLING : *La Vérité et le Menteur. Les paradoxes sui-falsificateurs et la sémantique des langues naturelles.*
Zlatka GUENTCHEVA : *Temps et Aspect : l'exemple du bulgare contemporain.*
Robert NICOLAI : *Parentés linguistiques (à propos du songhay).*
Claire BLANCHE-BENVENISTE (dir.) : *Le Français parlé. Etudes de grammaire.*
Stéphane ROBERT : *Système verbal et Enonciation en wolof.*
Herman PARRET (dir.) : *Le Sens et ses hétérogénéités.*
Danièle DUBOIS (dir.) : *Sémantique et Cognition. Catégories, prototypes, typicalité.*
Alain LE MARÉCHAL : *Problèmes de sémantique et de syntaxe en palau.*
Jacqueline GUÉRON et Jean-Yves POLLOCK (dirs) : *Syntaxe comparée et Grammaire universelle.*
Jean STÉFANINI : *Linguistique et Langue française.* Préface de Jean-Claude Chevalier, choix de textes et index par Pascal Bonnefois.
Jorge GIACOBBE : *Acquisition d'une langue étrangère. Cognition et interaction.*
Bernard LAKS, Annie RIALLAND (dirs) : *Architecture des représentations phonologiques.*
Michel LAUNEY : *Une grammaire omniprédicative. Essai de morphosyntaxe du nahuatl classique.*
Jean STÉFANINI : *Histoire de la grammaire.* Préface de Sylvain Auroux, choix de textes par Véronique Xatard.

SCIENCES DU LANGAGE

LANGUES INDO-EUROPÉENNES

Sous la direction de
Françoise BADER

CNRS EDITIONS

1994

Liste des collaborateurs

Françoise BADER, Ecole Pratique des Hautes Etudes, Paris, France

Dominique BRIQUEL, Université de Dijon, France

Claude BRIXHE, Université de Nancy, France

Enrico CAMPANILE, Université de Pise, Italie

Shaban DEMIRAJ, Université de Tirana, Albanie

Lambert ISEBAERT, Universités Catholiques de Namur et de Louvain, Belgique

Jay JASANOFF, Cornell University, New York, Etats-Unis

Charles de LAMBERTERIE, Ecole Pratique des Hautes Etudes, Paris, France

Manfred MAYRHOFER, Université de Vienne, Autriche

Craig MELCHERT, North-Carolina University, Etats-Unis

Anna PANAYOTOU en collaboration avec Brixhe

Heinz-Dieter POHL, Université de Klagenfurt, Autriche

Marie-José REICHLER-BEGUELIN, Universités de Fribourg et de Neuchâtel, Suisse

ISSN 0991-5877 — ISBN 2-271-05043-X

SOMMAIRE

INTRODUCTION

Françoise BADER

Ce livre est destiné à présenter à un public cultivé, mais non spécialiste, la plupart des langues indo-européennes, c'est-à-dire d'une famille qui a compris en Asie le tokharien, l'indo-iranien, l'anatolien, l'arménien, le phrygien, et en Europe le grec, le macédonien, le thrace, l'albanais, le slave, le baltique, le germanique, le latin et les autres langues d'Italie, le celtique, langues que nous ne connaissons que dans la mesure où elles ont été écrites, à des époques qui varient de la première moitié du second millénaire av. J.-C. pour les langues d'Anatolie au XVe siècle pour l'albanais.

1. PROBLÈMES DE CULTURE MATÉRIELLE

« Indo-européen » est un nom conventionnel qui ne s'applique à des hommes que dans la mesure où ils ont parlé ces langues, et bien d'autres dont nous entrevoyons parfois l'existence, mais dont nous ignorons parfois jusqu'à celle-ci quand elles n'ont pas été écrites. De leur culture, nous ne

connaissons que ce que nous pouvons appréhender par ces langues, leur vocabulaire, leur poésie et leurs mythes, qui, comme les langues, peuvent être étudiés par la méthode comparative, et à travers eux certains de leurs modes de pensée : ni lieu (« berceau ») ni date de genèse, sauf, précisément, ce que leur culture matérielle permet d'apercevoir.

1.1. Instruments de connaissance

1.1.1. Archéologie

Pour ce qui est de cette culture, deux ordres de faits peuvent fonder nos connaissances : l'archéologie, peut-être ; le vocabulaire, sûrement.

On associe souvent les hommes de langue indo-européenne aux « kourganes », nom russe des « tumulus » (d'une racine i.e. qui a fourni entre autres le nom de la « tombe ») qu'on connaît dans la période – 5000/ – 2000 par les cultures de la mer Noire et des régions de la Basse-Volga et du Sud du Caucase ; cette hypothèse, qui ne peut être étayée par aucun document épigraphique contemporain de ces sépultures, antérieures à l'emploi de l'écriture, reçoit un commencement de preuve du nom des « *tumulus* » (racine « être gonflé [et fort] »), ainsi que de certains textes qui les décrivent pour des hommes que nous savons par ailleurs avoir été de langues indo-européennes (et qui sont en général des chefs) : rois lydiens, comme Alyatte père de Crésus, pour le tumulus duquel l'archéologie confirme les dires d'Hérodote (1,93) ; rois scythes, dont les sépultures ont été localisées également par l'archéologie là même où les situait Hérodote (4,127), et contenaient des restes conformes à la description que donne l'historien des funérailles de ces rois (4,71) ; dans la Grèce homérique, Patrocle, dont les funérailles sont comparables à celles des rois scythes, et auquel les Achéens devront élever une tombe « large et haute » (*Iliade*, 23, 247).

1.1.2. Vocabulaire technique

Mais c'est la langue qui aide à cerner de plus près, et de manière plus vaste, la culture matérielle des hommes de langue indo-européenne dans le temps et, dans une moindre mesure, dans l'espace, grâce au vocabulaire : sont particulièrement précieux ici les termes techniques, dans la mesure où les techniques sont datables et localisables. La racine **ak-* (*h_2*ek-*), qui a donné par exemple des noms du « marteau » (type allemand *Hammer*) et de l'« enclume » (type grec ἄκμων), et s'est appliquée aux techniques

de la pierre taillée et polie, permet de saisir ces hommes à l'un des âges de la pierre, ce qui n'est pas susceptible de fournir le moindre élément de datation ni de localisation puisque les âges lithiques s'étendent sur des centaines de milliers d'années (à titre d'exemple, la « Pebble Culture » de l'*homo habilis* du Tanganyika pourrait dater de 1 750 000 ans). Il n'en va pas de même pour les termes de l'économie néolithique, caractéristiques du vocabulaire qu'atteint la reconstruction, par exemple noms d'animaux domestiques (chien, porc, ovins, bovins, cheval), et racines appliquées aux « semailles », au « labourage » et au « pâturage ». Ces termes sont particulièrement importants, car il est probable que l'économie néolithique a été le *primum mouens* de l'expansion des hommes de langue indo-européenne : l'explosion démographique qui en est la conséquence, et qui était incompatible avec l'économie de chasse et de cueillette, a poussé ces hommes à chercher de nouvelles terres de nourriture. Or nous savons à peu près où et quand cette économie a commencé : dans le Proche-Orient (plus précisément la région de l'actuel Irak) vers le IX/VIII[e] millénaire av. J.-C. ; cela nous fournit le lieu et le terminus *post quem* non pas de la genèse qui reste hors d'atteinte, mais de l'expansion de ces hommes.

1.2. « Berceau » et techniques néolithiques

C'est pourquoi le problème du « berceau », tant débattu, est en partie un faux problème : le développement des langues indo-européennes a dû durer des dizaines, pour ne pas dire des centaines de millénaires, prenant source dans le développement du langage lui-même ; et nous ne saisissons les hommes qui les ont parlées que sur la crête de la vague en quelque sorte : au début de cette croissance, dans une région qui, en tout état de cause, ne peut être le grand Nord circumpolaire de la fin de l'époque glacière où certains ont voulu les localiser au mépris des conditions de leur accroissement, mais qui peut varier selon qu'on se place aux débuts du néolithique, ou quelques siècles plus tard, quand celui-ci a porté ses fruits : la localisation dans l'espace et le temps du « berceau » des hommes de langue indo-européenne pose les mêmes problèmes de méthode que la reconstruction linguistique, nous amenant à tenir le plus grand compte, en chronologie relative, d'une successivité de couches dont bien souvent la plus récente est la seule à nous être accessible ; dans cette mesure, c'est vers le milieu du V[e] millénaire av. J.-C., dans la région des kourganes, que nous pouvons saisir les ancêtres des hommes qui ont migré vers l'Asie et l'Europe.

1.3. Migrations et toponymes

De ces migrations, nous apercevons quelques modalités, au plan de la langue, grâce, entre autres, à des toponymes et des ethniques.

On donnera pour exemples des premiers des noms propres qui entrent dans des formulaires poétiques. Ceux-ci nous font entrevoir quels termes ces hommes assignaient à leurs migrations : des terres bonnes pour la culture et la pâture, au pied d'acropoles naturelles fournissant le support de fortifications aptes à défendre les richesses qu'ils étaient venus faire fructifier dans leur quête de nourriture, et des terres bien pourvues de l'eau dont tant de mythes, par leurs taxinomies des variétés d'eaux douces, rappellent le caractère indispensable à toute fondation : πίειραν ἄρουραν (*Iliade* 18, 541), « terre arable grasse », avec un adjectif qui a donné le nom de la macédonienne *Piérie* et de l'*Ir-lande* (irl. *Ériu*, gall. *Ywerddon*); terres munies de « prairies » (anatol. **welu-> wilu-*, cf. *welu-<*welnu-*), comme Ἴλιον, anatol. *Wilu-sa-, -siya-*. Dans les formulaires, les qualités requises pour l'établissement sont distribuées entre le nom propre et l'adjectif : louv. *alati Wilusati* gr. Ἴλιον αἰπεινήν (*Iliade* 15, 558) « Les Hauts Praz escarpés », Ἴλιον... εὐκτίμενον πτολίεθρον (*Iliade* 4, 33), « Les Praz aux bonnes fortifications »; Λάρισσαν ἐριβώλακα (*Iliade* 2, 841), « La Rocheuse aux mottes soulevées », Καλυδῶνα... πετρήεσσαν (*Iliade* 2, 640), « Aiguebelle la rocheuse », αἰπεινῇ Καλυδῶνι (*Iliade* 14, 116) « Aiguebelle l'escarpée » (deux types de qualifications dissociées par exemple dans l'*Hymne à Apollon*, 41, où Samos est qualifiée de ὑδρηλή, et où les αἰπεινὰ κάρηνα, « les crêtes escarpées », sont celles de Mycale); etc.

Quant aux ethniques, ils seront étudiés à part (chap. III).

2. Problèmes de documentation

Les données dont nous disposons pour connaître la langue de ces conquérants ne constituent, encore, que la crête de la vague, car elles sont tributaires de l'écriture, en deux phénomènes relatifs à celle-ci : l'époque où les hommes de langue indo-européenne ont appris à écrire; et l'usage qu'ils ont consciemment voulu faire de l'écriture.

2.1. Langues anciennement documentées et exemple grec

La première est variable : il n'est pas fortuit que les plus anciennement documentées des langues indo-européennes appartiennent à une zone géographique où les recherches en matière d'écriture ont été les plus précoces. Il s'agit, d'une part, des langues anatoliennes (depuis –1700 environ), écrites dans le cunéiforme probablement d'origine sumérienne, qui a servi à noter des langues ni indo-européennes ni sémitiques, comme le sumérien lui-même ainsi que l'élamite, le hourrite, l'ourartéen, des langues sémitiques comme l'akkadien et l'assyrien, des langues indo-européennes, anatoliennes (hittite, louvite, palaïte) ou non (vieux perse) ; d'autre part, du grec.

Celui-ci a été noté en deux types d'écritures. Les unes, linéaire B qui note du grec mycénien, syllabaire chypriote, qui note un dialecte grec, sont syllabiques. Le linéaire B est employé aux XVe/XIIIe siècle av. J.-C. (datations qui sont l'objet de discussions) en des endroits divers, Péloponnèse (Pylos ; Mycènes, Tirynthe), Béotie (Thèbes), Crète, dont les documents mycéniens peuvent être plus anciens que ceux qu'a livrés le sol du continent, ici encore non par hasard ; car la Crète est, avec Chypre, comme un creuset où se sont formées diverses écritures (on ne peut, en l'état actuel des déchiffrements, les mettre en rapport avec la description linguistique qu'en donne le poète de l'*Odyssée*, 19, 172-6 : « la Crète, aux hommes innombrables, aux quatre-vingt-dix villes dont les langues se mêlent, 175, ἄλλη δ'ἄλλων γλῶσσα μεμιγμένη : s'y trouvent des Achéens, des Étéocrétois..., des Cydoniens, des Doriens..., des Pélasges ») : sinon celle qui a été imprimée sur le disque de Phaistos, difficile à dater et dont la provenance crétoise a été discutée, du moins une écriture hiéroglyphique essentiellement sur sceaux, probablement de la première moitié du premier millénaire, non déchiffrée ; et les deux linéaires qui peuvent en être des formes cursives, et sont apparentés entre eux, linéaire A (XVIIe/XVe siècle av. J.-C. ?), non déchiffré ; linéaire B. À ceux-ci est apparentée l'une des écritures attestées à Chypre, le syllabaire chypriote de Paphos qui note un dialecte grec, depuis le XIe siècle av. J.-C. en l'état actuel de nos connaissances. L'autre écriture est l'alphabet que des Grecs ont emprunté aux Phéniciens au plus tard vers –800, et qu'ils ont transmis à une grande partie du monde vers la fin du VIIIe siècle av. J.-C. par l'intermédiaire des Étrusques, dans la région de Cumes. Ceux-ci devaient apprendre à écrire à toute l'Italie, notamment aux Latins qui diffusèrent l'alphabet avec leurs conquêtes.

2.2. Écriture et idéologie de la mémoire

Quant à l'usage que les hommes de langue indo-européenne ont désiré faire de l'écriture, il est probable qu'il a pesé sur le caractère de notre documentation, essentiellement épigraphique à époque archaïque, et non pas littéraire. La raison de ce phénomène peut tenir à une idéologie de la mémoire, à double face.

2.2.1. « Écrire » et inscriptions

D'une part, la racine la plus ancienne pour « écrire », **pei-k-* (v.sl. *pĭsati*, etc.), a signifié « entailler » et a d'abord été employée pour « inscrire », et il en est de même pour nombre de ses renouvellements lexicaux comme γράφω. C'est que le premier usage de l'écriture a impliqué des matériaux impérissables, pierre puis parfois métaux, destinés à pérenniser d'abord un nom, nom d'un mort ou d'un dédicant dans les textes les plus nombreux qui sont des épitaphes et des dédicaces, parfois nom de celui qu'on voue à une malédiction, puis des textes promis à une longue vie, comme les lois ou les calendriers, etc. ; le montrent par exemple le formulaire des épitaphes irlandaises en écriture ogamique (attestée depuis le IV[e] siècle av. J.-C.), où le nom du défunt est au génitif dépendant de *ainm* « nom » exprimé ou implicite, ainsi que le rattachement du nom *ogam* à la racine de lat. *aiō*, *axāre* « nomināre ». Cet emploi de l'écriture est dû à une idéologie de la mémoire liée en grande partie à la mort, puisqu'elle a pour objet de fonder ce que les Grecs appellent, d'un tour formulaire hérité, le κλέος ἄφθιτον, le renom qui permet au nom de survivre en la mémoire des hommes.

2.2.2. Poésie et transmission orale

L'autre face de cette idéologie de la mémoire est l'exercice de celle-ci pour la transmission des œuvres poétiques, orales. Il faut en effet probablement étendre aux autres provinces du monde indo-européen ce que César dit des druides, *B.G.* 6, 14 : ils ne confient pas « à l'écriture la matière de leur enseignement, alors que pour tout le reste en général, pour les comptes publics et privés, ils se servent de l'alphabet grec. Ils me paraissent avoir établi cet usage pour deux raisons, parce qu'ils ne veulent pas que leur doctrine soit divulguée, ni que, d'autre part, leurs élèves, se fiant à l'écriture, négligent leur mémoire » (traduction Constans). Encore au Moyen Âge, les poètes traditionnels irlandais et gallois devaient mémoriser pendant de longues années un grand nombre de poèmes et de

contes, dont la transmission orale coexistait parallèlement avec la tradition manuscrite (coutume qui survit chez des conteurs populaires modernes d'Irlande, Écosse ou Bretagne). Auparavant, les Grecs, encore eux, étaient passés par là.

2.2.3. Homère et l'écriture

La plus ancienne des œuvres littéraires grecques à nous avoir été transmise par l'écriture montre qu'à cet égard une révolution s'est produite à l'époque d'Homère. L'*Iliade*, qui a derrière elle une longue histoire orale (dont témoignent en mycénien les noms *Hector* et *Achille* donnés à un berger de Knossos et à un théodule de Pylos), a été fixée par l'écriture peu après l'introduction de l'alphabet, probablement dans la seconde moitié du VIII[e] siècle av. J.-C. Il est probable que le poète a pris lui-même part à cette fixation, comme le montrent son nom et ses rapports à l'écriture. En forgeant son nom Ὅμηρος « qui agence (-ηρος) des compositions par liage » (*sh_2-om-o-*) d'épisodes qui ont pu lui être transmis oralement, et indépendamment les uns des autres, Homère a pu se référer à cette fixation, opérée par « liage » au moyen de techniques comme la répétition de formules en composition circulaire, etc. En tout cas, il a réfléchi à l'écriture dans les rapports que l'alphabet, introduit depuis peu en Grèce, pouvait entretenir avec le système phonologique de la langue grecque, qu'il connaissait forcément en tant que poète jouant d'allitérations. D'une part, s'il ne mentionne qu'une fois l'écriture, c'est dans un épisode (celui de Bellérophon, en VI) à l'intérieur duquel il crypte un tableau phonologique du grec ; d'autre part, en une seconde étape, il crypte un alphabet en fonction du système phonologique du grec dans la succession des quatre passages hermétiques connus sous le nom de « langue des dieux ».

Il n'est pas sûr que, ce faisant, il ne soit pas l'héritier d'une tradition relative aux rapports du poète avec l'écriture. D'une part, il se comporte comme l'un de ces druides (enseignant en vers) dont César nous dit qu'ils connaissent un usage officiel de l'écriture, *publicis privatisque rationibus*, et transmettent oralement leur enseignement parce qu'il est destiné à des initiés (*neque in vulgum disciplinam, efferri velint*). D'autre part, maître de la poésie, il l'est de l'alphabet et a en cela des correspondants ailleurs : en Islande, Óđinn est dieu de la poésie et des runes, dont le nom signifie « secret » (d'une racine qui, munie d'un élargissement [*h_1r-u-> *ruh$_1$-> *rū-/ *h_1r-eu-, etc.] ou non, s'est appliquée à la parole non transmissible avec diverses acceptions, la rumeur, lat. *rūmor*, le hurlement, skr. *ruváti*, *ráuti*, l'interrogation, gr. ἐρέω, ἐρευνάω tiré d'un nom *h_1r-eu-+-nā* à suffixe identique à celui du nom des runes) ; en Irlande, *ogam*, nom d'une écriture qui a fini par désigner tout langage cryptique, est étymologiquement relié à *Ogmios* (*$h_2$og-s-mio-*, de la racine de la parole transmissible), dont

Lucien comprenait le nom comme Λόγος (en accord avec l'étymologie par lat. *aiō* proposée), en racontant (cf. *Her.* 1) que le dieu traînait derrière lui des hommes attachés par les oreilles au moyen d'entraves d'or et d'ambre : n'y a-t-il pas là une représentation de la parole poétique qui « lie » par une magie à laquelle un Ulysse, après avoir bouché les oreilles de ses compagnons, résiste en se liant lui-même pour ne pas succomber au charme des Sirènes ? Quoi qu'il en soit, à partir d'Homère est neutralisée l'opposition entre sources épigraphiques et littéraires qui affecte notre documentation archaïque : en témoigne le fait que l'une des plus anciennes inscriptions grecques alphabétiques, datant du dernier quart du VIIIe siècle av. J.-C., donc d'une génération après la fixation par l'écriture de l'*Iliade*, et trouvée en Grande Grèce à Pithécuse, est celle de la « coupe de Nestor », qui se réfère en vers à un épisode de l'*Iliade*, et n'est pas, comme la totalité de nos inscriptions, une inscription appartenant à un dialecte (ionien-attique ; arcadien, chypriote, pamphylien ; éolien : lesbien, thessalien, béotien ; dorien et grec occidental).

3. Lacunes du livre et réflexions sur le latin

Parmi les langues indo-européennes dont il est parlé dans cet ouvrage ne figurent ni le grec ni le latin : j'ai choisi de les sacrifier par manque de place, espérant que mes lecteurs avaient par leur culture et/ou par leurs études plus d'occasions de les approcher que les autres langues. Du grec, il ne sera rien dit d'autre que ce qui vient d'être évoqué à propos des écritures, auquel pourrait être ajouté ce qui concerne les Écritures : le grec est la langue de la Septante dans laquelle a été transmise l'Ancien Testament au IIIe siècle av. J.-C. ; et c'est la langue de rédaction des Évangiles. Le latin, lui, est attesté, comme entre autres le grec, d'abord par des inscriptions (depuis le VIIe siècle av. J.-C., alors que la littérature latine, créée par des poètes tout nourris de l'hellénisme qui influença culturellement Rome de manière si profonde, commence dans le dernier tiers du IIIe siècle av. J.-C., au témoignage de ce qui nous a été transmis, c'est-à-dire un siècle après le début de la période alexandrine du grec). Mais on ne le prendra ici que comme point de départ de réflexions sur les modalités de l'indo-européanisation (après avoir pris le grec comme point de départ de réflexions sur la nature de notre documentation). Le latin fait partie de trois ensembles linguistiques ; dans deux d'entre eux (« indo-européen occidental » ; « italo-celtique ») il pose les mêmes problèmes que les autres langues italiques ; dans le troisième, il n'est que l'une de ces dernières.

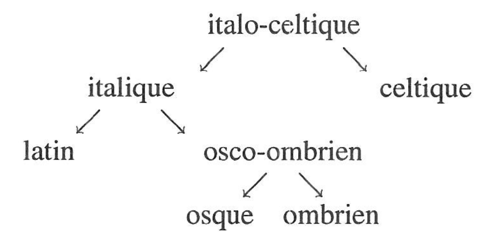
3.1. Indo-européen occidental

C'est «le vocabulaire du Nord-Ouest» qui définit l'indo-européen occidental; ce dernier comprenait, à l'époque où Meillet le conçut (cours professé au Collège de France en 1906-1907), le slave, le baltique, le germanique, le celtique (à l'époque essentiellement représenté par le celtique insulaire), l'italique (à l'époque représenté par le latin et l'osco-ombrien, sans le vénète, ni le messapien, d'identification bien postérieure). Cet ensemble est défini par des isoglosses lexicales (comme le nom de la «pomme», type angl. *apple*, cf. *Abella* «La (ville) des pommes», toponyme osque qualifié de *mālifera* par Virgile). De telles isoglosses peuvent résulter de phénomènes de contact, et non déterminer une parenté particulière, car seule la grammaire permet de déterminer une parenté linguistique (on sait par exemple que les Tsiganes, dont le nom figure peut-être chez Hérodote sous la forme Σίγυνναι (5,9), ont renouvelé leur vocabulaire au cours de leurs voyages, mais ont conservé assez de traits d'origine de leur langue pour qu'on puisse identifier celle-ci comme étant issue de l'indo-aryen du Nord-Ouest dont elle s'est séparée vers le V[e] siècle ap. J.-C.).

3.2. «Italo-celtique»

D'un autre poids à l'égard de la parenté linguistique sont les isoglosses que l'italique (au sens où l'entendait Meillet) présente avec le celtique; mais elles posent d'autres problèmes. D'une part, certaines des isoglosses communes à ces deux langues occidentales se retrouvent à l'autre extrémité, orientale, de la périphérie indo-européenne; certaines sont lexicales, comme celles qu'offrent le vocabulaire religieux de ces deux langues et celui de l'indo-iranien, d'autres grammaticales, comme celles qui ont des correspondants en hittite (par exemple des particules de phrase celtiques et anatoliennes) et/ou en tokharien (organisation du verbe latin en deux thèmes verbaux; intégration du parfait à l'un des deux, présent et/ou prétérit; emploi au présent médio-passif de particules en **-r(i)*, cf. lat. *datur*, hitt. *dattari*, en regard de la particule en **-i* de δίδοται, ou de skr. *dádate*); mais ces isoglosses ne permettent pas de définir une parenté particulière, car les seules qui soient aptes à le faire proviennent d'innovations, et non, comme ici, de conservations : ainsi que l'enseignait Meillet, les rameaux périphériques sont ceux qui se sont détachés le plus tôt du tronc commun. D'autre part, l'image des «rameaux» et du «tronc» qui vient d'être employée à dessein ne doit pas faire croire à l'existence d'un arbre généalogique tel que celui de Schleicher et de la première école

romantique allemande, qui aurait eu les provignements suivants :



Mis à part l'indo-iranien, on ne peut plus croire à l'existence de groupements tels que l'italo-celtique ; celui-ci en particulier a éclaté sous l'effet de deux types de facteurs : le déchiffrement du tokharien et du hittite, langues qui ont montré que certaines des innovations par lesquelles on avait cru pouvoir caractériser ce groupement n'étaient autres que des archaïsmes périphériques ; l'accroissement des données celtiques et italiques.

3.3. Latin et langues indo-européennes d'Italie

Car le latin, qui à l'aube de son histoire écrite n'est saisi lui-même que sous forme de patois (comme à Satricum où, au Ve siècle av. J.-C., le génitif est non en **-ī* mais en **-osio*, à Préneste, Ardea, Lavinium, Véies, etc.), fait partie d'un ensemble de langues indo-européennes d'Italie complexe, avec bien d'autres : falisque, à l'origine parler latin ; vénète, plus proche du latin que de toute autre langue italique, mais avec des isoglosses germaniques compréhensibles en raison de la proximité géographique du domaine germanique (et parfois un adstrat celtique consécutif aux invasions gauloises) ; osque et ombrien, et leurs parlers intermédiaires, plus proches les uns de l'ombrien, comme ceux des Volsques, Marses, Eques, Herniques, les autres de l'osque, comme ceux des Vestiniens, Marrucins, Péligniens au Nord, des Frentani, Samnites, Hirpini, Campani au centre, des Lucani, Bruttii, Mamertini au Sud ; préosque du Sud de l'Italie, antérieur à l'arrivée des Samnites (vers −530/−520), dont il existe deux groupes de chacun deux documents, sans qu'on sache s'ils appartiennent à la même langue : inscriptions de Nocera et de Vico Equense, pour l'instant isolées en pays salertinain, et les deux inscriptions de Casteluccio, dans le Bruttium ; messapien, qui diffère des autres langues d'Italie par la confusion de **ă* et **ŏ* en *a*, mais a en commun avec le latin un génitif en **-ī*, et présente un archaïsme périphérique notable pour « je suis » : *no*, préverbe en emploi prédicatif, qui apparaît à l'autre bout du domaine indo-européen, dans le verbe « être » tokh. B *nes-*/ A *nas-*<* *no-*+h_1s- « être » ; sicule, dont on ne sait s'il a des affinités avec le préosque du Bruttium, et qui est l'une

des trois langues de la Sicile ancienne (avec le sicane et l'élyme, non indo-européen).

La diversité de ces langues d'Italie fait qu'on ne peut plus croire, non plus, à l'existence d'un ancêtre commun au latin et aux autres langues. C'est à peine même qu'on entrevoit un « sabin » qui a pu être commun à l'osque et à l'ombrien : les plus anciennes inscriptions de ce groupe ont été trouvées en Sabine, Sabine tibérine pour la gourde de pèlerin de Poggio Sommavilla, du VIIe siècle av. J.-C., Sabine intérieure pour l'inscription de Cures, qui appartient du point de vue linguistique au « sud-picénien » ; or parmi les inscriptions de ce dernier, qui, outre celle de Cures, ont été trouvées surtout dans le Picenum, mais aussi chez les Praetuttii, Vestini, etc. (et qui datent probablement des VIe/Ve siècle av. J.-C.), l'une, le cippe de Penna Sant' Andrea, montre que cette population se considérait comme une *safina toúta*, « totalité politique de ceux qui ont fait leur territoire de conquête », fédération de ceux qu'unissait la langue par-delà des divergences dialectales assez minimes (de l'ordre des patois) pour que tous se comprennent ; et si, à l'époque où nous la saisissons, cette langue est ombroïde, un lien entre elle et l'osque se dessine à travers le territoire de langue osque qu'est le Samnium, contigu à la Sabine, et par la désignation *Sabellī* des peuplades aux parlers intermédiaires entre l'osque et l'ombrien.

4. MODALITÉS DE DIFFÉRENCIATION DES LANGUES INDO-EUROPÉENNES

Si le détail des faits qui rendrait compte des complications de la carte linguistique de l'Italie nous échappe, c'est par une conséquence des conditions de l'expansion des hommes de langue indo-européenne, qui ont abouti à un inextricable mélange d'héritages et de traits formés *in situ* : les héritages appartiennent eux-mêmes à des couches chronologiques diverses, si cette expansion s'est faite, non par des invasions qu'il serait anachronique de concevoir à l'image des grandes invasions du Moyen Âge, ne serait-ce que pour des raisons démographiques, mais par des infiltrations, chaque génération apportant un état de langue un peu différent de celui de la génération précédente, et se superposant en un lieu donné à celle-ci, dont la langue avait déjà un peu évolué ; quant au développement *in situ*, il peut être marqué tantôt par l'isolement des éléments d'une communauté, qui peut en accroître les particularités dialectales (par exemple en arcadien)

ou favoriser leur apparition (ainsi en vieil anglais), tantôt par des contacts (par exemple en pamphylien qui, coupé du monde grec, a subi fortement l'influence du substrat micrasiatique), tantôt par l'arrivée de populations de même langue et de dialectes différents en un même lieu (par exemple en Crète, où le mycénien a pu laisser des traces dans les traits non doriens du dialecte crétois).

C'est pourquoi il doit y avoir une part de vérité dans chacune des deux théories par lesquelles on a essayé d'expliquer la parenté des langues indo-européennes (cf. ici même, p. 59-60) : celle de l'arbre généalogique, sans laquelle on ne pourrait pas expliquer les archaïsmes périphériques ; celles des ondes, imaginée par J. Schmidt, ondes de départ, mais aussi ondes d'arrivée : ce serait un leurre que de concevoir la parenté de ces langues sous le seul angle génétique, sans phénomènes aréaux. En d'autres termes, la géographie a dû jouer un rôle important dans la formation des langues que nous connaissons comme apparentées par l'histoire : par exemple la langue des Vénètes est définie là où elle a laissé des traces écrites, comme linguistiquement apparentée aux langues géographiquement proches : gaulois en Bretagne armoricaine ; italique en Vénétie ; nous pouvons en inférer que la langue des Wendes, nom donné à leurs voisins slaves par les Germains à une date qui ne saurait être antérieure au Ier siècle av. J.-C., comme le montre la mutation consonantique **t>d* du nom, a dû être différente, sans pouvoir préciser si elle a été slave, ou autre ; et nous ignorons si la langue de ceux qui ont laissé leur nom au lac de Constance, le lac « vénète » qui s'est aussi appelé *Brigantinus* (> *Bregenz*) d'un nom celtique, a eu quelque chose à voir avec le celtique ; etc.

5. Exemples de langues peu accessibles à la connaissance

Il y a d'autre absents, mineurs au regard du grec et du latin, mais qui nous font entrevoir l'étendue de notre ignorance en matière de langues indo-européennes : ce sont les langues dont nous ne connaissons l'existence que par des sources indirectes, assorties parfois d'un nombre de textes inférieur à celui des doigts d'une main. On ne retiendra guère ici le « vieil européen », qui est un concept moderne, au contraire des langues des Pélasges, des Illyriens ou des Ligures, dont nous parlent les Anciens.

5.1. « Vieil européen »

La notion de « vieil européen » a été fondée essentiellement sur l'existence de noms de fleuves susceptibles d'avoir une étymologie indo-européenne. Mais il est impossible de définir linguistiquement la couche d'hommes qui les ont laissés. En voici quelques exemples.

D'une part, les noms de la *Drave* et de la *Save*, voisins, se ressemblent, mais posent des problèmes différents : de *sh_2-eu- « se déverser naturellement », le vocalisme est *a quelle que soit la langue ; mais de *$dr\text{-}^{e}/_{o}w$- « courir », un *a ne peut s'expliquer que par *o, selon un phonétisme qui ne pourrait être que germanique dans cette région en l'état actuel de nos connaissances ; le nom du « *Niemen* », proprement « courbure » (de *nem- « (se) courber »), est apparenté à celui de la peuplade celtique qui, à l'époque de César, est installée dans la « courbe » de la Loire, les *Namnetes* ; mais nous ne pouvons pas dire s'il est à l'origine lituanien, comme *Nẽmunas*, ou autre ; enfin (*last, no least*), le nom des *Aestii*, que Tacite, *Germ.*45, donne aux Baltes (et qui est devenu celui des Estes Finnois, par un transfert comparable à celui du nom germanique des Wendes aux Slaves), est dérivé de celui du fleuve *Aistà*, qui pose les mêmes problèmes phonétiques que le nom de la Save : il est tiré d'une racine à *a quelle que soit la langue (*$h_2$ei-s- « se mettre en mouvement par à coups », dit notamment des flots, comme la forme apparentée *$h_2$ei-g-, cf. Αἰγὸς ποταμός « Fleuve Bondissant » [et non « de la chèvre »], Αἰγύ-πτιος « Nil » « qui vole [court] par à-coups », Αἰγαί nom d'une ville de Macédoine connue pour les débordements du fleuve qui la traverse, et également nommée Ἔδεσσα [du radical *h_1w-ed- du nom de l'« eau », cf. ὕδωρ], etc.) ; nous ne pouvons donc pas dire si *Aestii* est germanique, baltique, ou autre, par exemple celtique, puisque selon Tacite la langue de ces hommes était « *Britannicae propior* ».

5.2. Pélasge grec

Autres sont les problèmes posés par le pélasge grec ; on n'en prendra pour exemple que l'onomastique qui figure dans le Catalogue des vaisseaux homérique, *Iliade* 2, 840-843 où les Pélasges sont donnés comme « habitants de la plantureuse Larissa... qui [ont] à leur tête... les deux fils du Pélasge Lêthos, fils de Teutamos » ; on ne retiendra que ce dernier et le toponyme (souvent lié par ailleurs à la mention de Pélasges), en laissant de côté les noms, grecs, de ces deux chefs, Hippothoos et Pylaios. La finale de Λάρισσα/ Λάρῖσα, comme celle de Παρνασσός/Παρνησός, qui

ont des correspondants en anatolien, ^{URU}*Larissa*, ^{URU}*Parnassa*, a joué un grand rôle dans la théorie du préhellénique, à tort comme nous le savons maintenant, grâce aux progrès faits en matière de phonétique historique des laryngales : une finale comme celle de ces dérivés d'appartenance-détermination, complexe, en *-ih_2-(*-eh_2-) [suffixes d'appartenance] +*-so-/-sā [morphème de détermination] est susceptible de deux traitements : assimilation de *-h_2s- > *-ss- ; chute de *h_2 et allongement compensatoire (cf. lat. *Ocrīsius, -ārius*, en *-ih_2-/*-eh_2-+*-si-o-). *Larissa* n'a alors rien de non grec (même dans son radical, qui peut être apparenté au nom de la « pierre » λᾶας), ni de préhellénique, le nom des Pélasges (ch. III, § 2) non plus ; seul est remarquable *Teutamos*, mais pas par son caractère non indo-européen : il a un correspondant exact en gaulois, et est un bon dérivé du nom de la « totalité politique » (cf. ici même, p. 82-83) ; le nom ne sort alors du grec que dans la mesure où celui-ci est démuni de *teutā dans son vocabulaire : des hommes de langue indo-européenne appelés Pélasges en Grèce, le passage homérique ne nous suggère rien d'autre que leur cheminement vers l'ouest dans la mesure où par ailleurs *teutā est occidental (si on exclut le rapprochement, parfois fait, de hitt. *tuzzi*- « armée »). Le pélasgique de Grèce ne témoigne donc de rien d'autre que de l'existence d'une couche linguistique non entièrement grecque, mais assurément indo-européenne (pour le nom des Pélasges, cf. ici même, p. 66-67).

5.3. Illyrien

L'illyrien, qui pose encore d'autres problèmes, nous renseigne, comme ce pélasgique homérique, sur l'importance de l'organisation politique de cette *teutā, qui, si nous avons eu raison de l'interpréter comme « totalité » transcendant l'opposition du soi et de l'autre, des conquérants et des conquis, doit jouer un rôle important dans la façon dont les hommes de langue indo-européenne ont remporté tant de victoires au moins linguistiques, en imposant leurs langues à une si grande partie du monde. Mais on est loin du pan-illyrisme de la première moitié de ce siècle. L'illyrien ne nous est plus connu que par des sources indirectes, qui nous donnent essentiellement des noms propres, puisque les sources directes sur lesquelles on en a pendant longtemps fondé l'existence ont disparu du fait de l'identification du vénète et du messapien comme langues indo-européennes d'Italie à part entière ; alors évanoui, l'illyrien ne peut plus être considéré comme l'ancêtre de l'albanais comme on propose souvent de le croire. Nous n'en possédons plus que des noms propres, dont l'étymologie fait souvent difficulté. Mais l'un, qui fait de l'illyrien disparu une langue indo-européenne, est du meilleur aloi : c'est le nom de la dernière des

reines illyriennes, vaincue par Rome au moment de la première guerre illyrienne en -229, *Teuta* (Plb. 2, 4, 6, etc.), qui a un doublet *Teuta-na* (Flor. 1, 21). Or ce dernier est de même formation que l'appellatif lat. *rēg-ī-na* « la femme du roi », le nom propre hitt. *Hassu-sar-na* « la femme du roi », « Rēgīna » ; et il peut être une transposition du vieux nom de la « reine » au plan de la « souveraine de la totalité politique ».

5.4. Ligure

Quant aux Ligures, leur aire a dû être plus étendue qu'elle ne l'est à époque historique : par exemple Caton, *ap.* Serv., *En.* 11, 700, dit qu'ils ont occupé une partie de l'Italie et de la Celtique ; mais il faut aussi abandonner le pan-ligurisme du siècle passé, qui reconstruisait sur des bases onomastiques (par exemple les formes à suffixe *-*sko*-) une aire ligure ayant recouvert toute l'Europe occidentale. La documentation épigraphique du ligure est misérable, une fois reconnu le caractère celtique du lépontique longtemps considéré comme ligure. Dans la vallée de la Magra (dite Lunigiana) qui aboutit à l'un des cols par lesquels s'établit une communication entre la région côtière ligure et la plaine du Pô, on a découvert une trentaine de « statues-menhirs » du premier âge du fer, sur lesquelles sont parfois représentées des armes de types connus par ailleurs comme celtiques. Des trois statues inscrites, une seule porte un texte lisible, MEZUNEMUŠUS, qui ne doit pas être gaulois, comme on l'a proposé à une époque où le lépontique était considéré comme ligure (en comprenant **Medionemōssos* « au milieu du bois ») : il n'y a pas de Gaulois dans cette région à l'époque que l'écriture assigne au texte (pas plus bas que −500) ; mais on ne peut en préciser la parenté. Il est vrai que la parenté du ligure et du celtique a pour elle la mention Κελτολίγυες des Anciens, et certaines gloses et noms propres, comme *Bergomum*, ligure d'après Caton (*ap.* Plin. 3, 124), ou *Comberanea*, nom d'une rivière près de Gênes, cf. gall. *cymmer*, bret. *kemper* « confluent » ; cependant, si on attribue au ligure le nom de fleuve *Porco-bera* « qui porte (comme [*Com*]*beranea*) des poissons » (quelle qu'en soit l'espèce, des noms de suidés ayant été souvent transférés à des poissons), ce nom a un *p*- celtique, mais qui pourrait appartenir à une langue d'Italie : ni le latin, qui aurait *Porci-(fera)*, par lequel Pline glose le nom ligure, ni l'osco-ombrien, qui aurait **porco-fera*, mais par exemple le vénète ; peut-être est-il possible de tirer parti ici de l'association faite par les Anciens entre le roi ligure Cycnos (« Cygne ») ami de Phaéthon, et le mythe de ce dernier, pleuré par ses sœurs dont les larmes devinrent des gouttes d'ambre, une indication de la proximité des Ligures avec les Vénètes de l'Adriatique auquel ce mythe

est lié. En tout cas, le ligure fait penser à une langue différente de celles que nous connaissons, qui aurait des isoglosses les unes celtiques, d'autres vénètes (et ayant pu se superposer à une (des) population(s) de langue(s) non indo-européenne(s)), isoglosses dont il n'est pas exclu qu'elles relèvent de phénomènes de contact).

5.5. Étrusque

S'il y a donc des absents dans ce livre, l'étrusque y a été retenu, bien que sa parenté indo-européenne ne soit pas établie. C'est qu'il offre des traits au moins typologiquement comparables à leurs correspondants indo-européens (ordre des mots de la phrase; existence de flexions nominales à génitifs en **-s* et **-l* homonymes de génitifs de langues indo-européennes; homonymies isofonctionnelles dans l'emploi de certains thèmes pronominaux; etc.); et les Étrusques venaient d'ailleurs, de Lydie selon Hérodote, dont la thèse est confirmée dans une certaine mesure par le caractère orientalisant de leur art. Tout en rappelant que le vocabulaire étrusque n'a rien d'indo-européen, mais que le vocabulaire importe peu dans l'établissement des parentés linguistiques, on tiendra cette langue pour non indo-européenne jusqu'à plus ample informé; mais le lydien, rangé par É. Benveniste dans l'édition des *Langues du Monde* de 1952 (I, 208-210) parmi les langues méditerranéennes, prend bien place maintenant, par suite de progrès de l'herméneutique, dans les langues anatoliennes; et, pour le même linguiste, « peut-être doit-on admettre que le peuplement de l'Étrurie s'est fait progressivement par des apports venus d'Asie Mineure et par fusion des nouveaux venus avec les autochthones » (I, 214), selon le mélange d'héritages et de développements *in situ* par lequel nous avons cru pouvoir expliquer de manière générale l'indo-européanisation linguistique : on a peut-être ici seulement une différence de degré entre les deux sortes de composantes.

Je remercie tous ceux qui ont accepté de collaborer à ce livre. Leurs contributions, toutes très savantes, diffèrent parfois par leur conception, d'où peut-être un certain disparate de l'une à l'autre. Mais j'ai été profondément soucieuse de respecter leur liberté, cette liberté dont le nom est parfois apparenté à celui des Aryens, hitt. *arawa-*, lit. *arvas*, v.irl. *aire* « libre », lexicalisation pronominale comme d'autres dans le domaine des

ethniques, et dans celui de la liberté (all. *frei*, angl. *free*, cf. skr. *priyá-* ; sl. *svobodĭ*). [1]

Annexe : Les termes « indo-germanique » et « indo-européen »

Le terme « indo-germanique » a été popularisé par J. KLAPROTH (*Asia polygotta*, Paris 1823). Voir G. MEYER, « Von wem stammt Bezeichnung “Indogermanen” ? » I.F. 2, 1893, p. 121-130.

Mais déjà en 1810 on trouve chez Conrad MALTE-BRUN, *Précis de la géographie universelle*, t. II (Paris : Buisson), livre 45 : « Nous nommerons en premier lieu la famille des langues indo-germaniques, qui règnent depuis les bords du Gange jusqu'aux rivages de l'Islande » (Traduction anglaise en 1882 : « Indo-Germanic »)

Le terme « indo-européen » se trouve d'abord en anglais : « Indo-European », chez Thomas YOUNG, c.r. de ADELUNG, *Mithridates oder allgemeine Sprachenkunde*, *Quaterly Review* 10, 1813, p. 250-292 (p. 255, 256, 264, 270, 273, 281).

Voir :

BUCK, C.D., (1904), « Indo-European » or « Indo-Germanic » ? *Classical Review* 18, p. 399-401

NORMAN, F., (1929) « Indo-European » and « Indo-Germanic ». MLR 24, p. 313-321 ; et pour d'autres désignations

KLAPROTH, Julius, (1810), « Indisch-Medisch-Slawisch-Germanisch », *Archiv für asiatische Literatur, Geschichter und Sprachkunde*, p. 81.

SCHLEGEL, (1919), « Indisch-lateinisch-persisch-germanische Sprachfamilie », *Wiener Jahrbücher der Literatur* 8, p. 413-468.

Je remercie Pierre SWIGGERS de m'avoir fourni ces indications. Et voir ici même, p. 49, note 3.

F.B.

1. Je donne plus de détails dans un article destiné à une Histoire de l'Europe, sous presse chez Einaudi à Milan : « Le problème indo-européen : langues locales vs. non locales en Europe entre – 600 et + 500 ».

CHAPITRE PREMIER

RECONSTRUCTION CULTURELLE ET RECONSTRUCTION LINGUISTIQUE

Enrico CAMPANILE

Alors que la reconstruction linguistique dans le domaine de l'indo-européen a comme objectif la connaissance de la grammaire et du lexique indo-européens, la reconstruction culturelle vise, elle, la culture des hommes de langue indo-européenne, dans chacun de ses éléments constitutifs, ainsi que dans ses structures sous-jacentes.

En partant du principe qu'une telle reconstruction est possible – et il sera question de cela plus loin –, celle-ci devra tenir compte de certaines considérations préliminaires, dont la première concerne les apports éventuels qui peuvent provenir de la recherche archéologique.

L'indo-européen est un concept avant tout linguistique, et ce n'est que par une implication naturelle qu'il nous faut convenir qu'il était parlé par des hommes, et que ces hommes possédaient aussi une culture (dont la langue elle-même était l'un des éléments constitutifs de premier ordre). Mais entre une langue et une culture matérielle – qui s'exprime dans les objets mis au jour et étudiés par l'archéologie – il n'existe aucun rapport de nécessité; il n'existe, au contraire, qu'un rapport historique et factuel, si bien que l'une ne nous permet pas de prévoir l'autre. Et

de plus, ce n'est que par le biais d'hypothèses que l'archéologue peut passer de l'analyse de la culture matérielle à des affirmations sur la culture non matérielle; plus ces hypothèses sont générales et éloignées de la matérialité de l'objet et moins elles apparaissent fondées. Quand, par exemple, Gimbutas (1980) caractérise la culture pré-indo-européenne de l'Europe comme « sédentaire, centrée sur la mère, pacifique, aimant l'art et munie d'un panthéon matriarcal de déesses en rapport avec la terre et l'eau », il est facile de comprendre qu'aucun de ces éléments ne part du témoignage immédiat du matériel archéologique, mais se situe au terme d'une série de déductions et de raisonnements qui pourraient cependant se développer dans d'autres directions (plutôt que « pacifiques », on pourrait tout aussi bien dire « munis d'armes extrêmement primitives » et qui n'apparaissent donc plus comme telles ou qui, en raison de leur nature physique, ne se sont pas conservées).

Il n'est donc pas facile de trouver un point commun entre la recherche linguistique et la recherche archéologique, même si l'on sait qu'il en existe un : il se trouve dans la société qui possédait cette langue et cette culture archéologique. Mais ce point commun ne peut être identifié que lorsque le linguiste et l'archéologue, indépendamment l'un de l'autre, se retrouvent au terme de leurs longues recherches au même lieu et au même moment. C'est le cas, par exemple, du monde gaulois préromain, où le linguiste identifie une langue (le gaulois) et l'archéologue une culture (la culture celtique), si bien qu'ensemble ils peuvent conclure que la société gauloise était caractérisée par ces deux éléments. Mais cela est possible dans le cas du gaulois parce qu'il peut être situé sans incertitude dans une dimension spatio-temporelle bien précise, grâce à la qualité de ses témoignages (citations, inscriptions, etc.), tandis que l'indo-européen, en tant que produit de reconstruction, ne permet pas de déterminations spatio-temporelles précises et ne peut donc jamais être superposé à un *facies* archéologique déterminé.

Le seul instrument qu'il nous reste pour une reconstruction culturelle est, en réalité, l'instrument linguistique; et la méthode, s'agissant là d'une culture qui n'est pas directement attestée, sera forcément de type comparatif et reconstructif. Le problème pourra donc se poser en ces termes : peut-on légitimement penser que la culture indo-européenne peut être appréhendée grâce à la comparaison et la reconstruction qui utilisent des matériaux linguistiques ? Ce n'est qu'après avoir répondu affirmativement à cette question que l'on pourra aborder d'autres questions plus techniques.

Au milieu du XIXe siècle, quand Kuhn instaura ce type de recherches, une réponse affirmative au problème fondamental que nous venons d'indiquer, était si évidente qu'on n'avait pas conscience de ce problème. Il semblait évident qu'un lexème – même un lexème reconstruit – comportât l'existence de la dénotation correspondante, et, par conséquent, le passage du « mot » à la « chose », dans le domaine de l'indo-européen

aussi, apparaissait absolument légitime. Au nombre croissant de mots indo-européens reconstruits correspondait ainsi un nombre égal de « choses » indo-européennes et il ne restait plus au chercheur qu'à les subdiviser et à les distribuer dans les cases appropriées : famille, agriculture, état, religion, etc. Paradoxalement, le nombre des « choses » finissait même par être supérieur à celui des mots, puisque par implication on pouvait aussi reconstruire des « choses » indo-européennes non attestées sur le plan lexical. Par exemple, non seulement on reconstruisait, à partir du lexème « père », l'existence de l'objet « père » dans la société indo-européenne (cette reconnaissance n'est pas une lapalissade, comme on pourrait le penser à première vue, puisqu'il y a aujourd'hui encore des sociétés « primitives » qui ignorent la fonction biologique du père), mais de la présence simultanée de « père », « mère », « fils », « frère », etc., on déduisait également l'existence de la famille, même si celle-ci ne peut être reconstruite comme lexème en indo-européen.

En réalité, le postulat même du parallélisme mots-choses est discutable et simpliste. En français, il existe des mots comme « chien », « lion » et « magicien », mais les dénotations correspondantes n'ont pas la même réalité dans la culture française actuelle. Les chiens existent en France, les lions n'y existent pas ; quant aux magiciens, ceux-ci n'y existent que de façon limitée et particulière, dans la littérature et dans les spectacles ou dans le langage métaphorique, mais non dans la réalité quotidienne, comme les chiens.

Pis encore, ce catalogue de « choses » déduites des lexèmes ne nous donne, en fait, pas plus que les lexèmes mêmes. Prenons le cas de indo-européen **rēĝs* « roi », correctement reconstruit à travers la comparaison de lat. *rēx*, gaul. *-rīx*, véd. *rāṭ*. On déduisait de cette forme qu'il existait un roi dans la société indo-européenne. Mais ce n'est qu'une étiquette que l'on récupère de cette façon – c'est-à-dire quelque chose qui ne diffère pas beaucoup d'un lexème –, puisque ce type de reconstruction culturelle ne nous dit pas – et n'est pas en mesure de nous dire – quelle réalité historique et fonctionnelle se cachait derrière ce titre. Qui était le roi indo-européen ? Était-ce une figure militaire, sacrée ou même économique, comme le voudrait Renfrew (1987) ? Quelles fonctions avait-il ? Dans quel cadre politico-social était-il inséré ? On ne peut même pas tenter de donner une réponse à des questions de ce type, tant que l'on demeure dans la séquence : comparaison lexicale → reconstruction lexicale → extrapolation existentielle.

La pauvreté intrinsèque de cette méthode lexicale – c'est-à-dire, d'une part, son atomisme, de l'autre, son incapacité de répondre à toute question d'ordre historique – fut perçue par un grand comparatiste qui avait médité à fond la leçon du structuralisme : É. Benveniste. Son *Vocabulaire des institutions indo-européennes* (1969) a, en effet, tenté de surmonter ces

limites, à travers un examen contextuel de tous les éléments pertinents d'un thème donné. Pour reprendre l'exemple du roi, Benveniste ne se limite pas, à la manière ingénue de Kuhn, à reconstruire le lexème indo-européen **rēĝs*, mais continue d'approfondir son sémantisme par l'intermédiaire de l'analyse étymologique. Celle-ci met en évidence le fait que **rēĝs* est l'ancien *nomen agentis* de la racine qui apparaît dans le lat. *regere*, et la nature sacrale du roi résulte du fait même que cette racine existait dans le domaine religieux, preuve en est la locution des augures romains *regere fines*. Ce caractère originel du roi est ensuite suivi par Benveniste dans ses continuations et variations historiques auprès de plusieurs peuples indo-européens, comme les Perses, les Grecs et les Romains. Enfin, certains mots grecs, tels que κραίνω « ordonner de faire », γέρας « part supplémentaire d'honneur accordée au roi » et τιμή « dignité permanente d'origine divine », sont examinés dans leurs références parallèles aux rois et au monde des dieux, ce qui donne une confirmation ultérieure du caractère sacré persistant de la monarchie indo-européenne en Grèce aussi.

Cette méthode, qui est aussi très suggestive, car elle met en jeu des données provenant des cultures les plus diverses et les discute avec la finesse et la compétence qui étaient propres à Benveniste, a, pourtant, deux limites profondes que nous pouvons mettre en évidence en réfléchissant précisément à ce qui a été dit à propos du roi.

L'auteur consacre à ce problème une cinquantaine de pages, mais, à y regarder de plus près, la thèse du caractère sacré de la monarchie indo-européenne repose entièrement sur un argument étymologique : **rēĝs* en tant que *nomen agentis* de la racine de (lat.) *regere* ; et, à son tour, la locution latine *regere fines* « tracer en lignes droites les frontières » (« C'est l'opération à laquelle procède le grand prêtre pour la construction d'un temple ou d'une ville et qui consiste à déterminer sur le terrain l'espace consacré : opération dont le caractère magique est visible », 1969, I, 14) garantirait le caractère sacré du nom. Mais cette argumentation est, en réalité, inacceptable : le matériau latin peut seulement prouver que *regere* était aussi un terme technique du langage religieux romain, mais rien ne nous autorise à étendre cette particularité à la racine indo-européenne correspondante ; cela exigerait au moins que la racine apparaisse avec la même fonction dans d'autres langues aussi, ce qui, manifestement, n'est pas le cas.

De ce point de vue, l'argumentation de Benveniste est plus pauvre que celle de Gonda (1956), qui, au moins, tentait de montrer que les notions d'« aide, protection » pouvaient être implicites dans la racine **reĝ-* « s'étendre, se tourner », si bien qu'en dernière analyse le roi serait celui qui tourne sa personne et sa force vers la défense de son peuple : le « protecteur ». La seconde limite de Benveniste, enfin, tient en ce que son analyse éclaire de manière adéquate des faits linguistiques et culturels déterminés d'époque historique, mais les projette dans l'époque indo-

européenne sur la base de leur parallélisme générique, sans considérer que la comparaison exige une vérification préliminaire de non-innovation, là où certains faits – comme les titres de la monarchie iranienne – sont manifestement des phénomènes novateurs (au sujet de ces titres, cf. Campanile 1974).

En réalité, la tentative d'enrichir la reconstruction lexicale traditionnelle par la théorie structuraliste des champs sémantiques et de l'accoupler à la recherche étymologique ne peut aboutir à des résultats convaincants, d'une part, parce que les champs sémantiques englobent des matériaux hétérogènes sur les plans étymologique et historique, de l'autre, parce que le sémantisme des racines indo-européennes est toujours si vague qu'il ne permet de trouver ni des signifiés rigoureusement délimités, ni encore moins des éléments connotatifs.

Il s'est donc formé, chez de nombreux chercheurs contemporains (Watkins, Meid, Schmitt, Bader, Campanile, etc.), quoique sans aucune unité de programmes et avec de nombreuses divergences méthodologiques, une double tendance : considérer la culture indo-européenne dans ses aspects intellectuels (religion, poésie, structure sociale, droit, idéologie, etc.) et tenter sa reconstruction avec une méthode autre que lexicale.

Le point de départ auquel tous se réfèrent plus ou moins directement est à situer dans les recherches de Dumézil : une masse énorme de livres et d'articles perpétuellement refondus et approfondis, sans que l'on puisse identifier parmi eux un texte canonique qui expose des méthodes et des résultats de forme définitive (une chaire de Philologie dumézilienne, aussi paradoxal que cela puisse paraître, serait moins inutile que tant d'autres chaires universitaires). Mais dans cette galaxie d'idées, d'intuitions et d'enquêtes, on trouve quelques points fermes. Le concept de tripartition est fondamental : il s'agit de la reconnaissance du fait que la culture indo-européenne (puis, de façon décroissante évidemment au fil du temps, celle de chacun des peuples indo-européens) analysait et créait toute forme de réalité en l'insérant dans l'une des trois catégories du sacré, du militaire et de l'économique. Ainsi, par exemple, quand on affirme la nature sacrée du roi, il ne s'agit pas là d'une définition générique parmi un nombre indéfini de définitions possibles, mais d'une insertion dans un schéma idéologique précis et limité, qui place automatiquement le roi en rapport avec toutes les autres réalités situées dans la catégorie du sacré, et en opposition avec toutes celles qui se trouvent dans les deux autres catégories.

Mais la position idéologique d'un élément culturel (il s'agit, répétons-le, d'un dieu, ou d'un mythe, ou d'une formule poétique, ou d'une donnée juridique, ou d'une institution, etc.) ne peut ressortir d'une simple donnée lexicale ; elle pourra ressortir, en revanche, d'une comparaison de la position que cet élément occupe dans chacun des schémas culturels : en d'autres termes, la sacralité du roi indo-européen est déductible non pas de son rapport avec le lat. *regere fīnēs* ou de l'analyse de la racine **reĝ-*, mais du

fait que dans les phases les plus anciennes des cultures indo-européennes, le roi apparaît de manière cohérente comme une figure éminemment sacrale et sacerdotale. Cette donnée ne résulte cependant pas de l'analyse des mots, mais de celle des textes : de fait, la méthodologie de Dumézil est déjà une reconstruction textuelle. C'est la voie que suivent aujourd'hui beaucoup de chercheurs, non seulement dans la classification par catégorie des éléments reconstruits (et une classification catégorielle, quelle qu'elle soit, est toujours nécessaire pour éviter l'atomisme reconstructif), mais, plus généralement, dans la reconstruction même de tous les éléments de la culture indo-européenne.

Du reste, dans une certaine mesure, Benveniste aussi (1932), peut-être sous l'influence de Dumézil, avait pratiqué une forme, quoique minime, de reconstruction textuelle quand, par exemple, il identifiait les trois classes originelles de la société indienne avec celles de la société iranienne, bien qu'il n'y ait entre elles aucune identité sur le plan du signifiant.

Quelquefois, naturellement, les reconstructions proposées par Benveniste peuvent effectivement être contestées, surtout pour sa tendance à aboutir à des conclusions trop détaillées et précises. Ce pourrait être le cas pour les « quatre cercles de l'appartenance sociale » (1969, I, 293 *sq.*), où l'affirmation selon laquelle la situation iranienne (*dam-* « famille », *vis-* « clan », *zantu-* « tribu », *dahyu-* « pays ») conserve fidèlement la situation indo-européenne n'a pas de base réelle au plan comparatif : à ce point de vue, il faut concéder quelque valeur à la critique de Zimmer (1987), tout en soulignant néanmoins que le point faible ne réside pas dans l'absence de correspondances lexicales, mais dans l'incapacité de superposer les structures sous-jacentes. A Rome, par exemple, on a la tripartition *domus/familia*, *gens*, *tribus*; nous savons par Homère que les Achéens étaient subdivisés en *phûla* et en *phrêtrai*, mais que les Doriens étaient *trikhai(w)ikes* « subdivisés en trois grands groupes » : ici, par ailleurs, le grec *wik-*, à la différence de l'av. *vis*, désigne la tribu (av. *zantu-*). Mais si Benveniste a été trop précis, les conclusions de Zimmer sur l'impossibilité de reconstruire la structure sociale indo-européenne sont probablement trop réductrices. Il faudrait plutôt relever la tendance commune à articuler la société en cercles d'extension croissante, et reconnaître qu'il ne s'agit pas là de phénomènes novateurs, puisque nous voyons, à partir des attestations les plus anciennes, que ceux-ci tendent à perdre progressivement de l'importance et de la signification. A Rome, par exemple, la tribu ne survit qu'en tant que collège électoral; à Athènes, la phratrie « remplit les fonctions du registre civil moderne » (Musti 1989, 93); le lit. *viešpats* ne désigne plus le chef du clan (comme l'av. *vispaiti*)-, mais équivaut simplement à « seigneur ». La conclusion la plus raisonnable, dans ce cas, semblerait être que la structure sociale indo-européenne ne peut être reconstruite dans le détail, mais que celle-ci existait et s'articulait en unités progressivement plus grandes, qui allaient vraisemblablement de la famille

(qu'il faudrait prendre, à mon avis, dans le sens de grande famille) à la totalité du territoire contrôlé par des personnes qui se sentaient unies par un lien commun.

A côté de ces critiques d'ordre factuel, qui portent atteinte aux résultats particuliers mais non à la méthode, il existe aussi des critiques d'ordre plus substantiel. Schlerath (1987), par exemple, estime que des difficultés et des apories dans la reconstruction textuelle naissent du fait que la conquête et la colonisation du monde pré-indo-européen furent l'œuvre d'une classe restreinte de guerriers, avec ses structures et ses valeurs, différentes de celles des Indo-européens dans leur ensemble; et il n'y a donc pas d'espoir de reconstruire deux cultures différentes avec un seul instrument linguistique.

En vérité, il est surprenant que ce soient précisément ceux qui nient toute possibilité de reconstruction culturelle qui le fassent en partant d'une affirmation d'ordre culturel-reconstructif; qu'est-ce d'autre, en effet, que la thèse selon laquelle les conquêtes indo-européennes auraient été l'œuvre d'une élite restreinte de guerriers, tandis que la grande masse de la population exerçait des activités plus pacifiques? Mais dans ce cas précis, en plus de la contradiction méthodologique, il est bon de souligner l'absence de fondement factuel de cette thèse. Il est vrai qu'à l'époque historique nous trouvons auprès de nombreux peuples indo-européens une classe restreinte de spécialistes de la guerre qui se consacrent intégralement à cette activité; mais en examinant les cas un à un, il est facile de voir qu'il s'agit là de faits novateurs qui, contrairement à ce que Wikander (1938) a été le premier à penser, ne peuvent être projetés jusque dans l'époque indo-européenne. Il suffira de quelques exemples pour s'en convaincre.

Les *fiana* irlandaises, qui correspondent bien à cette idée de bande de guerriers voués à la conquête et au pillage, sont attestées à une époque avancée du Moyen Âge; mais dans l'Antiquité, tous les historiens grecs ou romains s'accordent à nous présenter les invasions et les incursions celtes comme l'œuvre de peuples entiers, où tous les hommes valides combattaient; il n'est, du reste, pas pensable que de petites élites, même hautement spécialisées, aient pu vaincre de grandes armées aguerries comme celles des Grecs et des Romains. Il en va de même dans le monde germanique : le *berserkr* est lui aussi une figure du Moyen Âge, mais on a derrière lui le témoignage de Tacite (*Germanie* 13 *sq.*) d'après lequel tous les hommes sont tenus de porter les armes. Il en était de même à Rome où, à l'époque archaïque, la *gens* tout entière était mobilisée en temps de guerre (que l'on songe à la destruction totale des Fabii sur l'Allia). En réalité, l'idée même d'élite guerrière devra, à notre avis, céder la place à celle de peuple en armes; et il est significatif, sur le plan lexical, que des termes comme vieux-pers. *kāra-* et gal. *llu* désignent aussi bien la masse du

peuple que l'armée; il a dû se passer quelque chose de semblable autrefois en latin aussi, comme le suggère le rapport entre *populus* et *populari*.

Les réserves méthodologiques émises par Untermann (1985) apparaissent bien plus ponctuelles, même si nous devons maintenir une distinction claire entre ce qui se réfère à la reconstruction lexicale et ce qui concerne la reconstruction textuelle.

On fait à la première de ces méthodes des observations pertinentes sur l'impossibilité d'utiliser dans le domaine historique et culturel les résultats de la reconstruction lexicale. Quand nous reconstruisons, par exemple, « père » et « mère », rien ne nous permet de dire si ces termes avaient une connotation purement biologique ou s'ils comportaient l'existence d'une famille; et quand nous reconstruisons « cheval », nous ne savons rien sur son utilisation concrète, c'est-à-dire s'il était monté ou attelé à un char, ou bien simplement utilisé comme nourriture. Les propos d'Untermann confirment encore une fois l'inanité de la méthode lexicaliste.

L'exigence exprimée par Untermann de ne pas utiliser des données relatives aux contenus sans l'identité des signifiants nous paraît, au contraire, moins fondée; celle-ci reviendrait de fait à empêcher toute possibilité de recherche, du moment où toutes les langues ont plus ou moins fortement innové par rapport au lexique hérité. Cette exigence serait, en réalité, acceptable si nous voulions reconstruire le nom d'un élément culturel indo-européen; nous visons au contraire uniquement la reconstruction des contenus réels des éléments culturels.

Autres critiques : « La comparaison de concepts ou d'institutions sans un soutien étymologique n'a aucune force probatoire, dans la mesure où elle se réfère à des contenus qui peuvent ne pas être spécifiques à une langue ou à une famille linguistique; ces contenus peuvent à tout moment avoir été possibles également dans d'autres langues non indo-européennes dont les usagers avaient une culture voisine » (Schlerath, 1987). Et pourquoi donc tous les contenus de la culture indo-européenne devraient-ils être spécifiques à celle-ci uniquement? Cela nous rappelle de vieilles polémiques à l'encontre de Dumézil, où ses adversaires soutenaient que le tripartisme se trouvait aussi dans d'autres cultures (sémitique, chinoise, etc.), pensant ainsi porter un coup mortel à la reconstruction même de la culture indo-européenne opérée par Dumézil; et Dumézil eut le tort de défendre la singularité absolue du tripartisme indo-européen avec le courage de celui qui voit tous ses résultats scientifiques en péril, alors qu'il est évident que chaque culture possède très peu d'éléments absolument spécifiques et en partage beaucoup d'autres, à des degrés variables, avec d'autres cultures : après tout, nous faisons tous partie de la race humaine. Et si quelqu'un pensait identifier une culture uniquement par les éléments qui lui sont exclusifs, il risquerait, je le crains, de finir par ne définir la culture italienne que comme étant celle qui utilise les tomates dans toute sa nourriture à l'exception du café; ce qui serait peut-être déformer la réalité.

Et encore : « Toutes les reconstructions linguistiques sont toujours des abstractions (...) dont la position dans l'espace et le temps est largement approximative ; c'est pourquoi les formes concrètes d'existence des hommes qui parlaient cette langue que nous définissons comme "l'indo-européen" demeurent forcément inaccessibles » (Zimmer, 1987). Distinguons bien : l'indéterminabilité de l'espace géographique occupé par les Indo-Européens pèse certainement dans l'évaluation des formes concrètes d'existence, si nous donnons au terme ambigu « concret » le sens de « matériel » ; mais, comme nous l'avons déjà dit, l'intérêt majeur se porte aujourd'hui sur la culture non matérielle, que seul un déterminisme comme celui du XIXe siècle pourrait mettre rigoureusement en rapport avec les caractéristiques du milieu ambiant. Quant au temps, celui-ci ne peut être déterminé en valeurs absolues, mais peut l'être, et de manière rigoureuse, en termes relatifs : le comparatisme reconstruit, par définition, ce que l'indo-européen possédait au moment de sa dissolution. Que celle-ci ait eu lieu quatre ou sept millénaires avant Jésus-Christ, cela est tout à fait insignifiant pour nos problèmes.

Et encore : « Abstraction faite des doutes que l'on peut avoir sur le fait de reporter des situations sociales de sociétés (proto-)historiques à une préhistoire de deux ou trois mille ans plus ancienne, il serait étrange que l'organisation de base de la société soit demeurée inchangée en dépit de toutes les modifications politiques, économiques et culturelles. Entre l'époque de la communauté indo-européenne primitive, vraisemblablement presque exclusivement pastorale, et l'apparition des premières cultures organisées d'Anatolie et de Grèce, qui cultivaient la terre et pratiquaient le commerce, il s'est passé tant de choses qu'il nous faut presque exclure une conservation statique de l'ancien mode de vie » (Schlerath, 1987). Nous remarquerons ici aussi au passage que l'auteur utilise, pour nier la possibilité de la reconstruction culturelle, un élément (le pastoralisme des Indo-Européens les plus anciens) qui ne peut lui être suggéré précisément que par la reconstruction culturelle. Mais le point essentiel est que les développements de la civilisation ne doivent pas être vus sous une forme catastrophique et totalitaire, telle que chaque innovation détruit tout son passé ; il est plus conforme à l'expérience historique de reconnaître que de très nombreux éléments des cultures précédentes survivent aux bouleversements les plus voyants. Et le premier devoir évident du comparatiste consistera précisément à identifier les éléments novateurs qui ne doivent pas être utilisés dans la reconstruction (monnaie, commerce maritime, techniques militaires, etc.). Mais ce sont là des questions techniques qui ne concernent pas la méthodologie générale.

Nous avons dit que la reconstruction textuelle était une procédure qui consiste à comparer non pas des lexèmes, mais des données textuelles ; il s'agit donc, en substance, d'une comparaison de contenus. Prenons un exemple.

C'est un fait bien connu que, dans le domaine de la religion indo-européenne, les noms des dieux qui apparaissent dans les différents panthéons ne sont pas comparables entre eux pour permettre la reconstruction des *nomina sacra* indo-européens. La seule exception est constituée par véd. *Dyaus*, gr. *Zeus*, lat. *Iuppiter*, qui permet de reconstruire un indo-européen **Dyēus* (l'autre exemple que l'on proposait autrefois, à savoir la comparaison entre véd. *Varuṇa-* et gr. *Ouranós*, est phonétiquement insoutenable).

Laissons de côté, même s'il est important, le fait que la figure de Dyaus dans le panthéon védique est très différente de celle de Zeus et que cette dernière a fortement influencé celle du lat. Iuppiter, si bien que les fonctions et l'encadrement de **Dyēus* sont difficiles à définir. Cependant, il demeure surprenant que le panthéon indo-européen s'avère, nominativement, non reconstructible. L'explication la plus raisonnable de cette situation est qu'il y a eu d'importants renouvellements lexicaux, qui ont été rendus possibles par la structure tripartite dans laquelle s'articulaient les divinités : ce qui était important, ce n'était pas la personne du dieu (comme cela advient dans les religions modernes), mais la place qu'il occupait dans le schéma de l'idéologie indo-européenne et les fonctions qui en découlaient : le fait que les dieux (frères) de la troisième fonction s'appellent *Nāsatya* (Inde), *Castor* et *Pollux* (Grèce), *Freyr* et *Freyia* (Germains du Nord) ou *Romulus* et *Remus* (Rome ; cf. Campanile, 1988) n'avait donc aucune importance.

Cette fragmentation des noms divins ne devra cependant pas nous suggérer une attitude de renonciation ; elle justifie au contraire les tentatives de comparaison (et de reconstruction), même si elles partent de différents éléments onomastiques. Et l'instrument ne sera pas les noms (qui, par définition, sont différents les uns des autres), mais tout ce qui caractérise la figure divine au niveau textuel : épithètes, mythes, parenté, fonctions, cultes : de cette façon, nous pourrons reconstruire non pas le nom, mais la substance d'un dieu indo-européen. Nous donnerons ici un bref exemple, pour lequel nous devrons remonter à un point très éloigné : jusqu'à une sainte chrétienne.

Dans la tradition chrétienne irlandaise, saint Patrick et sainte Brigitte sont les figures de saints les plus éminentes. Mais alors que le premier est une figure historique sûre, même si sa chronologie est controversée, la seconde rentre sûrement dans la grande masse des saints qui n'ont jamais existé ; tous les chercheurs s'accordent aujourd'hui à dire qu'elle n'est rien d'autre que la transposition dans le panthéon chrétien d'une ancienne divinité païenne. Nous connaissons bien cette déesse préchrétienne, par l'intermédiaire de sources érudites irlandaises et de textes épigraphiques de Bretagne dans lesquels elle apparaît sous sa forme latinisée *Brigantia* ; à la base de la forme irlandaise et britannique, il y a, respectivement, un **Brigentī* et un **Brigantī*, qui remontent tous deux à celt. **Br̥gn̥tī*.

Un tel nom prête en soi déjà à réflexion, puisque c'est un adjectif (cf. véd. *bṛhatī*, av. *bərəzaiti*) signifiant « haute, noble » : il s'agit d'une épithète qui s'est transformée en nom propre dans le monde celtique. Les parallèles ne manquent pas : dans certaines localités de la Grèce, par exemple, les deux Dioscures sont appelés « les Rois » (*Wanake*), les pythagoriciens appelaient « Illustre »(*Euklês*) le dieu de l'outre-tombe, Athéna est souvent appelée « la Vierge »(*Parthénos*) ; nous ferons plus loin un parallèle plus pertinent encore.

Un premier indice sur la nature originelle de Brigitte nous est fourni par le Ṛgveda, où *bṛhatī* « haute, noble » n'est pas seulement une épithète de la déesse Aurore (véd. *Uṣas-*), mais va dans certains hymnes jusqu'à se substituer au nom : « Avant tout le monde s'est levé *Bṛhatī*, victorieuse, conquérante de la récompense » (R̥V 1, 123, 2). Il ne s'agit évidemment pas d'une preuve : ce n'est qu'une sollicitation. Voyons s'il y a d'autres éléments qui confirment l'équation celt. Brigitte = i.e. Aurore.

Dans la *Vie de sainte Brigitte* (une biographie du Moyen Age qui n'a aucune valeur documentaire, mais qui conserve fidèlement d'anciens éléments de la tradition préchrétienne), on répète avec une grande insistance que Brigitte naquit à l'aube et que sa mère la mit au monde sur le seuil de la maison, avec un pied à l'intérieur et l'autre au dehors : c'est-à-dire que l'enfant naquit à mi-chemin entre l'obscurité de la maison et la lumière du jour. Or, dans la poésie grecque, dès Homère, le qualificatif le plus typique de l'Aurore est « née à l'aube » (*ērigéneia*) ; d'autre part, l'emplacement intermédiaire de l'Aurore est précisé dans le monde védique, en la définissant comme celle qui « fait tourner les deux espaces contigus » (R̥V 7, 80, 1) : les deux espaces sont ceux de la lumière du jour et de la nuit par rapport auxquels l'Aurore sert pour ainsi dire d'axe. Il y a donc une parfaite correspondance entre une donnée biographique, une épithète et un théologème dans trois cultures différentes.

Une confirmation ultérieure du fait que Brigitte est l'Aurore nous vient de sa propre généalogie : son père s'appelait Noir, son grand-père Sombre : c'est l'Aurore qui naît des ténèbres de la nuit.

La nature lumineuse de l'Aurore est soulignée à maintes reprises dans les sources védiques : « Jeune fille rayonnante aux vêtements lumineux » (R̥V 1, 113, 7), « vêtue de lumière » (*ibid.* 1, 124, 3). La sainte chrétienne n'a pas de telles épithètes, mais dans sa biographie on trouve toujours un lien miraculeux avec la lumière, avec le feu, avec les rayons du soleil : une colonne lumineuse s'élève de sa tête, les rayons du soleil lui servent de portemanteau, le châle qu'elle a sur sa tête émet des flammes : c'est la traduction concrète de cette luminosité qui, dans la poésie védique, est exprimée par l'épithète *vibhātī* « resplendissante ».

La même racine *$*bheH_2-$* « resplendir » se retrouve en grec à propos de l'Aurore, mais sous une forme plutôt indirecte. Dans l'hymne homérique au Soleil, en effet, nous lisons : « Muse Calliope, fille de Zeus, commence

à célébrer le Soleil lumineux, qu'Euruphaassa aux yeux de jouvencelle (*boôpis*) engendra au fils de Gaia et du Ciel étoilé» (*H.H.* 31, 1 *sq.*). Qui est la mère du Soleil ? Dire que c'est l'Aurore n'est pas seulement une question de bon sens, mais est également conforme à ce que les hymnes védiques répètent à maintes reprises ; par exemple : «les Aurores resplendissantes ont donné naissance au Soleil» (7, 78, 1). Euruphaassa est donc l'Aurore. Nous nous trouvons encore une fois dans le cas où une épithète devient nom propre, puisqu'Euruphaassa «resplendissante» est étymologiquement et sémantiquement comparable à véd. *vibhātī*, épithète de l'Aurore. Le fait nouveau consiste en ce qu'Euruphaassa ne se substitue pas au nom de l'Aurore, mais en constitue un dédoublement, exactement comme cela advient pour *ērigéneia* «née à l'aube», autre épithète de l'Aurore qui apparaît chez Hésiode comme un doublet de l'Aurore : «l'Aurore engendra les vents Zéphyr, Borée et Notus (...), Erigéneia engendra Eosphore (l'étoile du matin !)» (*Théogonie* 378 *sq.* ; Campanile, 1987).

Revenons à Brigitte. Si elle est, comme cela semble évident, la transformation chrétienne de l'Aurore indo-européenne, qu'offre donc de nouveau et d'utile ce résultat du point de vue de la reconstruction de la culture indo-européenne ?

Le fait est que les chercheurs ont admis avec trop de facilité l'existence d'une telle déesse indo-européenne, en se fondant sur des argumentations substantiellement lexicalistes, sans prêter trop d'attention aux réalités factuelles et textuelles.

En fait, la seule culture où la déesse Aurore apparaît de manière évidente et incontestable est la culture védique. Mais les choses sont déjà bien moins claires dans la culture iranienne. Ici, en effet, l'Aurore (*Ušah-*) n'apparaît en tant que divinité que dans un texte extrêmement tardif (G. 5, 5) ; ailleurs, elle désigne l'aurore en tant que phénomène naturel ou comme point cardinal (orient). Par exemple, dans une strophe du 5e Yast (§ 61) étudiée par Benveniste, on lit : «A la fin de la troisième nuit, il atteignit l'aurore rayonnante et, vers l'aurore, il implora Ardvi Sura Anahita.» Il est évident ici que l'aurore indique simplement l'orient («il atteignit l'orient et, tourné vers l'orient, il implora», etc.) ; Benveniste l'écrit justement avec une minuscule. Mais nous reviendrons plus loin sur l'Aurore iranienne.

Existait-il en Grèce une déesse Aurore ? Il suffirait d'un regard sur Homère pour répondre par l'affirmative ; mais ce serait une réponse non fondée : l'Aurore grecque (Eos) est une figure purement littéraire. Particulièrement aimée par les poètes, elle est inconnue du peuple en tant que déesse. Elle n'a ni temple ni prêtre, ni cultes, ni fêtes ; son nom n'apparaît pas dans les indices épigraphiques. La comparaison entre gr. *Eos/Auos* et véd. *Uṣas-* est phonétiquement correcte, mais ne peut impliquer aussi que toutes deux aient été des divinités.

L'aurore romaine était encore moins une déesse. A raison, Radke (1965) se limite à la définir comme la « Wiedergabe der gr. Eos » ; il s'agit donc d'un personnage littéraire, étranger à toute réalité cultuelle.

Dans le cas de l'Aurore aussi, on a donc fait un saut dangereux du lexique à la réalité cultuelle ; et ce qui a joué un rôle certainement déterminant dans ce saut, c'est le fait que l'Aurore védique est effectivement une divinité : le préjugé selon lequel le védique serait plus ou moins l'indo-européen est toujours opérant dans le domaine de la recherche indo-européenne.

A ce point, si nous prenions simplement en considération l'Aurore védique, grecque et latine, nous n'aurions pas d'argument pour reconstruire une divinité indo-européenne ; et l'iranien ne ferait qu'ajouter des éléments d'incertitude, car l'attestation isolée et tardive de la déesse Aurore suggérerait plutôt l'hypothèse d'une innovation ou d'un emprunt qui ne pourrait pas, en tout cas, entrer dans un discours reconstructif.

C'est ici que réside l'importance de Brigitte : celle-ci ne nous donne pas une forme ultérieure du nom indo-européen de l'Aurore, mais, au-delà de son nom, elle nous prouve qu'autrefois la déesse Aurore existait aussi parmi les Celtes. Et l'équation véd. *Uṣas* = v.irl. *Brigit* nous permet de reconstruire non pas un nom, mais une divinité indo-européenne.

Mais il y a plus encore. Une fois établie l'existence de la déesse, nous pouvons à nouveau réfléchir sur les matériaux qui, à première vue, ne se prêtaient pas à la comparaison. Nous avons dit que, du point de vue documentaire, l'Aurore grecque est une figure littéraire. Mais l'essentiel est ailleurs et nous pouvons nous demander : figure littéraire créée *ex novo* par les poètes grecs, ou bien conservation littéraire d'un culte très ancien mais désormais en désuétude en Grèce dès les débuts de l'époque historique ? A notre avis, c'est cette dernière solution qu'il faut retenir ; ce qui nous le prouve, c'est le fait que l'Aurore, bien qu'elle soit un personnage littéraire, présente toute une série d'éléments (épithètes, connotations, etc.) qui sont parfaitement identiques à ceux de la déesse Aurore en Inde et en Irlande. Nous avons déjà vu que l'épithète *ērigéneia* trouve confirmation dans la biographie de Brigitte ; *Euruphaassa*, qui, avant d'être un doublet, avait dû être une épithète de l'Aurore, est immédiatement comparable à l'épithète védique *vibhātī* ; l'Aurore grecque a un char et des chevaux, exactement comme l'Aurore védique (qui a pour cette raison l'épithète *bṛhatratha-* « au char élevé ») ; l'Aurore grecque sous le nom d'Euruphaassa, est mère du Soleil, tout comme en védique ; l'Aurore grecque est mère de Memnon qui, selon la belle étymologie de Saussure (1881), est à comparer avec le véd. *manman-* « pensée, souvenir, poésie » et la déesse Brigitte pré-chrétienne était, d'après les érudits irlandais, déesse de la poésie.

Tout cela ne peut être une coïncidence et ne trouve sa justification que dans le fait que la tradition poétique grecque a conservé fidèlement, sur le plan littéraire, une déesse. Et c'est précisément cela qui nous autorise

à réintroduire dans la comparaison ce personnage. Comme nous le disions précédemment, à Rome, l'Aurore n'est qu'un simple reflet littéraire de la littéraire Aurore grecque. Mais c'est le mérite de Dumézil d'avoir démontré que cette figure divine survit aussi à Rome sous un autre nom. Il s'agit de la déesse Mater Matuta.

On observera avant tout qu'il s'agit là d'une divinité authentique, avec son temple (construit, d'après la tradition, par Servius Tullius et refait par Camille), avec ses fêtes (les Matralia, célébrées le 11 juin), avec ses miracles pour porter secours aux généraux romains. Or, il y a déjà, en faveur d'une identification de Mater Matuta avec l'ancienne Aurore, un indice d'ordre linguistique, souligné avec raison par Leumann (1977, § 296) : l'adjectif *mātūtīnus* se présente comme un dérivé de *Mātūta* ; mais comme *mātūtīnus* signifie « relatif à l'aube », il n'est pas déraisonnable de supposer que c'était là le sens le plus ancien de *Mātūta*. Mais à part cette étymologie synchronique, il y a un argument culturel, mis en relief par Dumézil, qui montre clairement comment Matuta est le nom qu'a pris à Rome l'Aurore indo-européenne.

Le jour des Matralia, en effet, les matrones fécondes et *univirae* se réunissent dans le temple de la déesse et en chassent rituellement et brutalement une esclave ; elles adressent ensuite à Matuta des prières non pas en faveur de leurs enfants, mais en faveur des enfants de leurs sœurs. Comme le remarquait Dumézil, tout cela présente d'impressionnantes analogies avec l'hymnologie védique de l'Aurore : elle chasse tout élément négatif et promeut le lever du soleil, qui, selon la conception védique, est non seulement le fils de l'Aurore, mais également son neveu en tant que fils aussi de la Nuit (Rātrī) : ce que l'Aurore fait en Inde (chasser toute négativité et favoriser le neveu), à Rome ce sont les pieuses matrones qui le font.

Revenons, à ce point, aux faits iraniens. Comme nous l'avons vu, ici la déesse Aurore n'apparaît que très tard, dans une prière ; auparavant, il n'existait qu'une aurore, pas une déesse. D'un point de vue historique, ce n'est pas surprenant, puisque la religion originelle de Zarathoustra refusait comme des démons les anciennes divinités indo-iraniennes (Mitra, Vāyu, Vr̥thragna, etc.) et celles-ci ne réapparaissent que dans un deuxième temps, à travers un syncrétisme. Mais cette déesse Aurore, même si elle n'est rappelée que dans une phrase très courte, possède encore un trait caractéristique de l'ancienne déesse : un char avec des chevaux rapides (*ravat̰.aspa* et *ranjat̰.aspa*). Du reste, il est significatif que l'aurore non déesse présente, elle aussi, exactement la même épithète (*viāuuaiti* « resplendissant ») qui, dans les hymnes védiques, est l'apanage exclusif de la déesse Aurore (*vibhātī*) et qui a dû exister aussi en Grèce (voir ce qui a été dit au sujet de Euruphaassa) ; et tout comme en védique *bhāsvatī* « lumineuse » est épithète de l'Aurore, en avestique le synonyme

de même radical *bāmyā* devient désignation de l'Aurore ; et cette épithète était ressentie comme tellement liée à l'idée d'Aurore qu'en pahlavi l'aube est appelée *ušbām* (litt. « aurore resplendissante »).

On peut donc supposer que la déesse Aurore a été tout d'abord effacée de la religion de Zarathoustra, mais que l'aurore (comme donnée physique) a conservé les épithètes qui étaient autrefois propres à la déesse ; et quand l'aurore entra, provenant d'on ne sait quelle tribu iranienne, dans le panthéon de Zarathoustra, elle amena avec elle des éléments très anciens fidèlement conservés, si bien qu'il y a une coïncidence substantielle entre la représentation laïque et la représentation religieuse de l'aurore.

La façon dont peut procéder une reconstruction textuelle doit être claire désormais. Pour les vieux chercheurs, l'aurore indo-européenne se fondait avant tout sur l'équation lexicale :

véd. *uṣas-* : av. *ušas-* : gr. *Eos/Auos* : lat. *Aurora* et, dès lors qu'elles étaient aussi des déesses, on concluait, à la fin, en l'existence d'une déesse indo-européenne Aurore.

Au lieu de cela, nous avons vu que la situation est bien plus complexe et qu'à côté de déesses authentiques il y a aussi des personnages littéraires que l'on ne peut accueillir immédiatement dans la comparaison, mais dont le genèse doit être examinée avec soin. Notre comparaison se fondera donc sur les matériaux suivants, qui sont d'ordre textuel et non pas lexical : véd. *Uṣas-* : v.irl (déesse païenne *Brigit* + sainte *Brigit*) : lat. *Mater Matuta* : gr. *Eos* (déesse disparue comme telle dès avant les documents les plus anciens mais conservée dans la tradition poétique + *Erigéneia* + *Euruphaassa*) : av. *ušas-* (sous son double aspect de divinité entrée tardivement dans le panthéon zarathoustrien et de phénomène naturel qui conserve des éléments propres de l'ancienne divinité).

L'existence d'une déesse indo-européenne de l'Aurore étant reconstruite sur cette base, il reste à éclaircir ses connotations, sa fonction, sa place dans l'idéologie tripartite.

En ce qui concerne les connotations, il suffit de se référer à ce qui a déjà été dit au cours de cette discussion. L'Aurore est conçue comme un être féminin, lumineux, qui se déplace sur un char, mère ou tante du Soleil (selon que celui-ci est considéré comme son fils ou comme le fils de sa sœur Nuit) ; elle est jeune et belle.

Sa fonction première est de donner lieu à la naissance du nouveau soleil, garantissant ainsi aux hommes une nouvelle journée de vie. Et c'est pour cela que dans un hymne védique elle est appelée « esprit de la vie » ; il s'agit de l'un de ses aspects qui n'est explicite que dans le monde indien, mais qui a dû être très important pour la foi populaire, puisqu'aucun autre dieu du panthéon indo-européen n'avait la fonction bénéfique d'assurer et de prolonger la vie des hommes. C'est précisément cet aspect d'« esprit de la vie » qui pourrait expliquer, selon notre hypothèse (Campanile, 1990), l'énorme importance assumée par sainte Brigitte en Irlande, où, dans

plusieurs textes (rédigés par des religieux érudits, remarquons-le bien), elle s'identifie et se substitue à la Madone même. Elle est également déesse de la poésie ; et cette fonction naît du fait que la première prière de la journée a lieu à l'aube. Comme le dit un hymne védique, l'Aurore est mère puissante de la prière, c'est la jeune femme qui attire avec l'inspiration et pousse les hommes au lieu du sacrifice (R̥V 5, 47, 1) ; et la prière, nous le savons bien, n'était pas une expression intime et informelle, mais une œuvre formalisée des poètes-prêtres. Dans le monde grec, il reste de cela un souvenir dans la tradition qui voit en l'Aurore la mère de Memnon, c'est-à-dire de la prière-poésie fondée sur la mémoire ; plus simplement, en Irlande, la Brigitte préchrétienne est mentionnée en tant que déesse de la poésie dans le glossaire de Cormac.

Quelle place avait l'Aurore dans le tripartitisme indo-européen ? Dumézil a montré qu'un petit nombre de divinités n'appartenaient pas spécifiquement à l'une des trois fonctions, non pas parce qu'elles sont étrangères à l'idéologie tripartite, mais parce qu'elles étaient fonctionnelles et préliminaires à chacune des trois fonctions ; on peut donc parler de divinités trifonctionnelles. Et c'est certainement le cas pour l'Aurore. Sa spécificité réside dans le fait de mettre un terme à la non-activité nocturne et de faire débuter les multiples activités diurnes ; elle est donc le préliminaire à toute activité sacrée, militaire ou économique. De même, elle est le début de la prière, quelle que soit la divinité à laquelle elle s'adresse : ici aussi, l'Aurore est fonctionnelle dans le culte de toute divinité de quelque fonction que ce soit.

La reconstruction culturelle a donc connu un développement prodigieux depuis l'époque lointaine où de grands précurseurs, comme Kuhn ou Pictet, se limitaient à comparer entre eux le véd. *Dyaus*, le gr. *Zeus* et le lat. *Iuppiter* pour reconstruire un indo-européen. **Dyēus* et en déduire, avec une argumentation étymologique (lat. *diēs*, etc.), qu'il s'agissait de la divinisation du ciel, quitte à se demander ensuite s'il s'agissait du ciel diurne et lumineux ou d'un ciel de pierre, dans la mesure où, dans plusieurs langues, le même lexème signifie tantôt « ciel », tantôt « pierre », et tantôt tous les deux.

On pourra reprocher à toute reconstruction culturelle des erreurs, des imprécisions, des ingénuités, des exagérations dans l'analyse des textes, des idées préconçues ; il est cependant difficile aujourd'hui de mettre en discussion la légitimité des tentatives visant à reconstruire avec des moyens linguistiques – les seuls dont nous disposions – la culture indo-européenne ; moins que jamais nous n'accepterons le principe que la linguistique indo-européenne doit avoir pour seul objectif la reconstruction de l'indo-européen. Si la linguistique indo-européenne – comme toute linguistique diachronique – est une science historique, alors sa légitimité consistera à faire de l'histoire : par conséquent, histoire de la langue, mais aussi histoire de la culture.

Bibliographie (arrêtée en 1990)

BENVENISTE É. (1931). – « Deux notes iraniennes », *BSLP* 32, p. 86 *sq.*

– (1932). – « Les classes sociales dans la tradition avestique », *JA* 221, p. 117 *sq.*

– (1969). – *Le Vocabulaire des institutions indo-européennes*, 2 vol., Paris.

CAMPANILE E. (1974). – « Ant. pers. *xšayaθiya xšayaθiyānām* », in *Studi Bolelli*, Pisa, p. 110 *sq.*

– (1987). – « Histoire et Préhistoire d'une formule poétique indo-européenne ». *EIE* 21/24 (= G. Dumézil in memoriam I), p. 17 *sq.*

– (1988). – « Tradizione storiografica romana e ideologia indoeuropea » (=*Alle origini di Roma*, a cura di E.C., Pisa, p. 9 *sq*).

– (1990). – *La Ricostruzione della cultura indoeuropea*, Pisa.

DELBRÜCK B. (1889). – *Die indogermanischen Verwandtschaftsnamen*, Leipzig.

DE SAUSSURE F. (1881). – ΑΓΑΜΕΜΝΟΝ, *MSLP* 4, p. 432 (=*Recueil*, p. 403).

DUMÉZIL G. (1974). – *La Religion romaine archaïque*, Paris.

GIMBUTAS M. (1980). – Introduction, *JIES* 8, p. 1 *sq.*

GONDA J. (1956). – « Semantisches zu idg.*rēĝ-* "König" und zur Wurzel *reĝ-* ("sich ausstrecken") », *KZ* 73, p. 151 (= *Selected Studies*, I, Leiden 1975, p. 415 *sq*).

LEUMANN M. (1977). – *Lateinische Laut- und Formenlehre*, 1.B. (Neuausgabe), München.

MUSTI D. (1989). – *Storia greca*, Bari.

RADKE G. (1965). – *Die Götter Altitaliens*, Münster.

RENFREW C. (1987). – *Archaeology and Language*, London.

SCHLERATH B. (1987). – « Können wir die indogermanische Sozialstruktur rekonstruieren ? Methodologische Erwägungen » *Studien zum indogermanischen Wortschatz*, hgg. von W. Meid, Innsbruck, p. 249 *sq.*

UNTERMANN J. (1985). – « Ursprache und historische Realität. Der Beitrag der Indogermanistik zu Fragen der Ethnogenese », *Studien zur Ethnogenese*, Opladen, p. 133 *sq.*

– (1889). – *Kratylos* 14, p. 45 *sq.*

WIKANDER St. (1938). – *Der arische Männerbund*, Lund.

ZIMMER S. (1987). – « Indogermanische Sozialstruktur ? Zu zwei Thesen Emile Benvenistes », *Studien zum indogermanischen Wortschatz* (cité), p. 315 *sq.*

CHAPITRE 2

LA MÉTHODE COMPARATIVE. PROBLÈMES ÉPISTÉMOLOGIQUES EN DIACHRONIE LINGUISTIQUE

Marie-José REICHLER-BÉGUELIN

1. PRATIQUES SPONTANÉES DE L'ÉTYMOLOGIE

1.1. Le cratylisme et ses procédures

Socrate à Hermogène : Ignores-tu, bienheureux homme, que les premiers noms établis ont été comme enfouis par ceux qui voulaient en rehausser la magnificence ? Ils y ont ajouté des lettres ou en ont ôté pour l'euphonie et ces mots ont été tordus dans tous les sens, soit avec l'intention de les embellir, soit par l'effet du temps. Témoin le mot *catoptron (miroir)* : ne trouves-tu pas étrange qu'on y ait intercalé un *r* ? Voilà ce que font, j'imagine, ceux qui, sans aucun souci de la vérité, se mêlent de modeler la prononciation : à force d'insérer des lettres dans les noms primitifs, ils finissent par les altérer au point que personne au monde ne saurait

comprendre ce que le mot peut bien signifier. (Platon, *Cratyle* 414 c, trad. É. Chambry)

Lorsque Socrate, en imitateur malicieux de la méthode de Cratyle, cherche à savoir pourquoi un *r* a été intercalé dans le mot grec *kátoptron* « miroir » (*Cratyle* 414c), il pose en fait implicitement la question du lien qui unit cette entité lexicale aux autres mots de la même famille, notamment à l'adjectif *kátoptos* « visible », et au futur *katópsomai* (qui laisse supposer un présent non attesté **kat-óptomai*; ces formes sont à l'évidence dépourvues de *r*). Mais il opère ce rapprochement sans procéder à l'analyse que nous proposerions aujourd'hui de la finale du mot *kátoptron*. Cette analyse y dégagerait le suffixe *-tron* très répandu de noms d'agent, par rapprochement de la famille paradigmatique du grec *phére-tron*, *spék-tron*, *béle-tron...*, cf. latin *-trum* dans *aratrum*, sanskrit *-tram* dans *vás-tram*, etc. ; la voie serait ainsi ouverte à la reconnaissance d'une racine verbale *op-* signifiant « voir », attestée dans d'autres mots grecs comme *ópo:pa* « j'ai vu » (d'autres rapprochements plus sophistiqués comme celui de hom. *ósse* « les yeux », ainsi que le témoignage des langues apparentées : lat. *oculus*, etc., conduiraient ultérieurement à postuler une forme plus ancienne de la racine $*ok^{w}$-< $*h_3ek^{w}$-).

Alors que le mot *kátoptron* reste inanalysé dans le *Cratyle*, pour *me:khané:* « ruse », autre terme examiné par Socrate, une décomposition nous est proposée. Mais elle est non grammaticale du point de vue moderne, car uniquement fondée sur le rapprochement, approximatif et purement fortuit à nos yeux, des signifiants de *me:khané:* (« ruse »), *mé:kos* (« longueur ») et *aneîn* (« accomplir ») :

> *Mèkhanè* me paraît indiquer *le fait d'aller loin dans l'accomplissement* d'un travail; car *mèkos* signifie une grande longueur. C'est donc de ces deux mots, *mèkos* et *anéïn (accomplir)*, qu'on a composé *mèkhanè*. (*Ibid.*, 415 b)

Une paraphrase, qui se veut étymologique, du sens de *me:khané:*, est ainsi avancée sur la base d'une ressemblance de sonorités, d'une vague analogie formelle qui sert d'étai à une spéculation d'ordre sémantique. On est tout proche des jeux de langue très populaires en français moderne comme la charade ou le rébus, ou de procédés mis en œuvre par la poésie ou la littérature, que j'illustrerai par une citation de Michel Leiris :

> Matador : d*ama*ssé il *mat*e l*a mort* et la *dor*e *d'aromat*es.

Cet énoncé définitoire exploite, comme dans l'exemple de *me:khané:*, diverses associations lexicales, lesquelles sont engendrées, par allitération et anagramme, à partir de la structure phonique du mot *matador* : j'ai souligné, dans la citation, les phonogrammes qui assurent une réitération du *definiendum* dans le *definiens* (cf., du même auteur, le titre même du recueil : « Glossaire j'y serre mes gloses »).

1.2. L'étymologie populaire

Les exemples rapportés au paragraphe précédent illustrent une pratique de l'étymologie bien ancrée dans l'usage spontané des sujets parlants : la recherche du « vrai » sens des mots, qui peut être accessoirement conçu comme leur sens ancien, primordial, est très souvent confiée aux rapprochements intuitifs qu'autorisent les analogies de signifiant. L'activité métalinguistique des jeunes enfants fourmille de phénomènes comparables ; l'enfant qui cherche à trouver une motivation à un mot quelconque la fonde très volontiers sur une similitude sonore, rime, assonance, paronomase, etc. :

> Madame *Sonia*, parce que c'est elle qui fait *sonner* la cloche » (S., 6 ans, citée par Bonnet et Tamine, 1985, je souligne)

C'est par un même processus également que les sujets parlants remotivent, au prix souvent d'une déformation, certaines unités lexicales soit d'origine étrangère, soit devenues inanalysables : fr. *choucroute*, adapté de l'all. *Sauer-kraut*, littéralement « fermentée-herbe », a ainsi été rapproché de *chou*, de manière purement interne au français, et au mépris du sens originel ; quant à a.fr. *coute-pointe*, de *coute* « couverture », variant de *couette*, et *pointe*, participe de *poindre*, « cousue », il a été transformé en *courte-pointe*, ce qui indique une réinterprétation totale, entérinée par l'histoire, d'un vocable que l'évolution linguistique avait rendu obscur (dans ce cas, un *r* a bel et bien été intercalé dans le mot... Pour d'autres faits relevant de ce qu'on appelle l'étymologie populaire, voir F. de Saussure, 1916 = 1972, 238-241).

1.3. Étymologie spontanée et « vérité » linguistique

Dans tous ces exemples, qui illustrent une pratique non théorisée de l'étymologie, notons que rien ne différencie méthodologiquement une explication de type synchronique, interne à un état de langue donné (*matador*, *Sonia*), d'une explication à prétention diachronique ou historique (*me:khané:*). Confrontées à une entité lexicale quelconque, l'une comme l'autre en exploitent sans règles strictes le matériau signifiant afin de produire une paraphrase interprétative. Qu'elle vise ou non une origine, cette paraphrase se veut éclairante, révélatrice d'une vérité profonde sur le mot (*étymos* = « vrai »). Mais, comme elle n'est pas poursuivie systématiquement, pour elle-même, elle a souvent les apparences d'une réinterprétation idiosyncrasique, mise au service d'une visée argumentative propre au dis-

cours où elle s'insère ; cela même quand elle prétend reposer sur un savoir linguistique :

> ... la parole littéraire est fatalement *abolie*, dans le sens étymologique du mot : c'est-à-dire : venue de loin (*aboleo*). (G. Agamben, *L'Origine et l'Oubli* ; en fait, lat. *aboleo* signifie « je détruis »)

Longtemps, l'activité étymologique a ainsi fonctionné par rapprochements ponctuels, sur la base de procédures heuristiques hasardeuses, inaptes à mettre un frein à quelque spéculation que ce fût. Sur de tels fondements, il a par exemple été possible de soutenir durablement l'idée, encore représentée au XVIIIe siècle dans l'article *Langue* de l'*Encyclopédie*, que tous les idiomes d'Europe descendaient de la langue biblique et adamique qu'était l'hébreu, les entités lexicales ayant subi dans les diverses langues des permutations de syllabes ou de « lettres », bien entendu postulées *ad hoc*[1]. D'un point de vue moderne, une telle hypothèse est infondée, aberrante : d'où la tentation de condamner purement et simplement l'étymologie spontanée comme un sous-produit de l'ignorance, une source potentielle de faux savoirs sur la langue.

Toutefois, le problème du statut qu'il convient d'accorder à certaines analyses que la linguistique dite scientifique ne saurait reprendre à son compte n'est pas aussi simple qu'il y paraît. Depuis Saussure, on sait en effet que la langue, conçue comme produit de l'activité des locuteurs, « fait des fautes, mais qui n'en sont pas » :

> La langue ne peut pas procéder comme le grammairien, elle <est à> un autre point de vue et les mêmes éléments ne lui sont pas donnés, elle fait ce <qui par> le grammairien est considéré comme des erreurs <mais> qui n'en <sont> pas, car il n'y a de sanctionné par la langue que ce qui est immédiatement reconnu par elle. (Saussure, *CLG*/Engler 2759-60, I R 2.65-66)

En tant que symptôme du comportement grammatical des sujets parlants, l'étymologie spontanée véhicule bel et bien une « vérité » sur la langue, même si ses hypothèses n'ont pas droit au label de certitudes historiques. C'est ainsi que l'analyse de *kátoptron* par intercalaison d'un *r*, à première vue invraisemblable, atteste indirectement une réalité linguistique de la synchronie attique : car les inscriptions de l'époque présentent souvent une variante *katropton* (avec anticipation du *r*), confirmant que la forme n'était plus ressentie comme un dérivé incluant le suffixe *-tron* ; l'existence de cette variante fournit par ailleurs une explication au brouillage morphologique dont le mot semble avoir été l'objet. Dépourvue de valeur historique probante, l'analyse socratique vaut donc néanmoins comme indice des

1. Sur de telles bases, l'idiome ancestral pouvait tout aussi bien être le flamand, le français, le suédois, le scythique, au gré des idéologies et selon les postulats de traditions parallèles : cf. Olender, 1989, 13-20.

réinterprétations auxquelles les sujets parlants soumettent constamment les unités de leur langue.

2. Comparaison et reconstruction dans la méthode historique

2.1. Conditions d'application de la méthode : l'analyse du mot en sous-unités

Il fallait, on s'en doute, un renouveau méthodologique important pour que l'étymologie acquît le statut d'une véritable science, visant une connaissance rationnelle des états linguistiques antérieurs aux premiers documents attestés, une connaissance élaborée pour elle-même et sans visée idéologique. Un des principes importants de la mutation théorique attendue se situe dans la prise en compte d'unités grammaticales, *morphologiques*, de rang inférieur au mot, mot dont l'analyse cesse d'être fondée sur un découpage d'ordre phonique ou syllabique, remotivé sémantiquement *a posteriori* et de façon plus ou moins arbitraire. Dans l'*Encyclopédie* encore, mais cette fois-ci sous l'entrée *Étymologie* (article dû à Turgot et inspiré des travaux de De Brosses), on rencontre l'ébauche d'une méthode nouvelle, où l'« invention » propre à cet art conjectural qu'est l'étymologie se trouve contrôlée par une « critique » permettant d'« apprécier la certitude des étymologies ». Selon l'encyclopédiste, l'activité étymologique doit notamment se fonder sur une décomposition grammaticale des mots, entités parfois complexes qui, on le sait, ne recoupent pas forcément les unités significatives minimales.

> L'examen attentif du mot dont en cherche l'*étymologie*, et de tout ce qu'il emprunte, si j'ose ainsi parler, de l'analogie propre de sa langue, est donc le premier pas à faire. [... ;] avant de chercher l'origine d'un mot dans une langue étrangère, il faut l'avoir décomposé, l'avoir dépouillé de toutes ses inflexions grammaticales, et réduit à ses éléments les plus simples.

A ceux qui expliquent lat. *(insula) britannica* par hébreu *baratanac* « pays de l'étain », il est ainsi objecté que le mot latin est un dérivé formé de deux éléments (*britann* + *ica*), qui doivent recevoir des explications distinctes. Ainsi, la spéculation génétique se trouve confrontée à une première exigence de rigueur : l'identification préalable des unités morphématiques (racines, affixes, désinences).

Quelques décennies plus tard, le principe du découpage morphologique recevra une impulsion déterminante avec la révélation des analyses produites sur le sanskrit par les grammairiens indiens, dont la technique descriptive dégageait avec une grande précision la racine et les autres sous-unités du mot (Rousseau, 1984, 312-316). Or, l'attention portée aux morphèmes grammaticaux (marques de genre, de nombre, de cas, de temps, de personne, etc.) et non plus seulement aux éléments lexicaux, aura une incidence de première importance sur la scientificité de la méthode en linguistique historique. En effet, les premiers, au contraire des seconds, sont peu susceptibles d'emprunts de langue à langue : leur coïncidence peut donc se révéler particulièrement décisive quand il s'agit de fonder sur des bases incontestables l'existence d'une communauté d'origine.

2.2. Comparer, reconstruire

2.2.1. Séries de correspondances et apparentement linguistique

La méthode comparative proprement dite naît donc avec la « découverte » du sanskrit, langue de l'Inde avec laquelle les Occidentaux étaient en contact depuis le XVIe siècle, mais dont les ressemblances nombreuses avec les langues de l'Europe (latin, grec, gotique, vieux-slave, langues baltes, celtique, etc.) ne frappèrent véritablement les esprits qu'à la fin du XVIIIe siècle et ne furent étudiées systématiquement qu'à partir du début du XIXe siècle, avec les travaux révolutionnaires de Bopp et de Rask [2]. Aux premiers comparatistes, on doit d'avoir substitué aux rapprochements ponctuels, isolés, la quête de *séries* de correspondances, permettant d'exclure la possibilité d'emprunts réciproques ou de coïncidences fortuites, et donc de conférer une assise solide à la notion d'apparentement linguistique. Dès lors, il fut admis que seule la mise en lumière de *réseaux de concordances* réguliers, comme ceux qui sont faciles à constater entre les langues romanes, toutes issues du latin, était de nature à apporter la preuve que deux langues remontent effectivement à un même prototype (cf. Meillet, 1925).

---

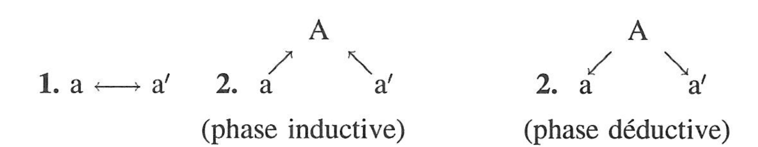
2. On ne présentera pas ici d'histoire de la grammaire comparée, la perspective de ce chapitre étant orientée vers les stratégies épistémologiques. Pour des renseignements complémentaires sur les premiers comparatistes, voir par exemple Meillet, 1937, 453-483 ; Leroy, 1971, 17-57 ; Mounin, 1974, 156-217 ; Auroux (éd.), sous presse.

2.2.2.1. Des formes observées au prototype : induction, déduction

> *...deux langues sont dites parentes quand elles résultent l'une et l'autre de deux évolutions différentes d'une même langue parlée antérieurement.* (Meillet, 1937, 16)

La simple constatation des relations entre langues d'une même famille, intéressante en soi, n'a toutefois pas suffi à combler l'ambition des comparatistes : depuis plus de 150 ans, la grammaire comparée des langues indo-européennes s'est donné pour tâche plus précise de *reconstruire* l'idiome hypothétique qu'on a convenu d'appeler « indo-européen » [3], c'est-à-dire la langue-source non attestée dont découlent, entre autres, la plupart des langues modernes d'Europe.

La méthode comparative comporte donc deux étapes : la *comparaison* proprement dite, fondée, comme on l'a vu, sur l'observation de correspondances régulières (phonétiques, grammaticales, lexicales, syntaxiques, voire phraséologiques) d'une langue à l'autre; puis la *reconstruction*, consistant à restituer, sur la base des correspondances établies, le prototype indo-européen d'où proviennent les diverses formes historiquement transmises. À l'intention de ses étudiants, F. de Saussure a limpidement caractérisé la démarche méthodologique de l'indo-européaniste (Reichler-Béguelin, 1980, 24) : dans un premier temps, une relation est établie entre deux observables a et a′ (étape 1 dans le schéma ci-dessous); dans un second temps (étape 2), un raisonnement inductif permet de poser A, la forme postulée dont on considérera que découlent respectivement a et a′ (phase déductive du raisonnement) :



Ainsi, la comparaison de lat. *medius*, skr. *mádhyaḥ*, gr. *mésos* (dial. *méssos*, *méttos*), got. *midjis*, etc., autorise la reconstruction d'un prototype **med^{h}jos* « situé au milieu », apte à rendre compte des diversités phonétiques que présentent les formes attestées. A noter que a et a′ peuvent aussi être des variantes propres à une seule et même langue : ainsi *ger-* et *ges-* dans lat. *gérō*, *gestus*, ramenées toutes deux par reconstruction interne à une forme plus ancienne unique **ges-*, dont la sifflante a subi le rhotacisme en position intervocalique, mais non devant la dentale *t* de *gestus*. La reconstruction des signifiants se fonde sur la reconnaissance de règles d'évolution

3. Sauf dans la tradition allemande, qui utilise encore l'adjectif *indogermanisch*, « indo-germain » : sur les avatars terminologiques, non innocents, qui marquèrent la naissance de la grammaire comparée, voir Olender, 1989, 27-30.

phonétique précises et propres à chaque idiome pour une période donnée, règles que les néo-grammairiens de la seconde moitié du XIX[e] siècle baptisèrent du terme de « lois » phonétiques.

2.2.2.2. Cas de figure de la reconstruction : unification (réduction à l'unité) *vs* distinction

Dans les exemples examinés précédemment, le raisonnement inductif conduisait à poser, à partir d'observables formellement distinctes, un prototype unitaire, présumé plus ancien et subsumant leurs divergences : il s'agissait, intellectuellement parlant, d'une démarche d'unification. Mais l'évolution linguistique n'est pas faite que de différenciations. Il arrive aussi que des unités initialement différentes aient été confondues, réduites à l'unité au cours du temps. On se trouve dès lors devant la tâche, cognitivement plus difficile, d'avoir à *distinguer*, derrière une entité attestée, les (au moins) deux éléments hétérogènes dont elle représente éventuellement le syncrétisme :

```
f. reconstruites :   A         B    (éventuellement C, D...)
                       ↖     ↗
f. attestée :             a
```

C'est grâce à une démarche du second type qu'autour des années 1875, la vision que l'on se faisait du système vocalique indo-européen s'est progressivement renouvelée. Jusqu'alors en effet, il était admis que le *a* indo-iranien, qui répond très souvent en latin et en grec tantôt à *e* et tantôt à *o* (cf. skr. *jánas-*, qui nivelle la diversité vocalique observable dans gr. ***génos***, lat. *genus* < **genos*), reflétait directement l'état de la langue primitive, les autres langues présentant des différenciations contextuelles mal expliquées (on opérait donc dans ce cas une reconstruction par unification). Les efforts conjugués d'Amelung, Brugmann, Collitz, Saussure et quelques autres, ont conduit à une inversion de perspective. Ces auteurs ont progressivement démontré que le ***a*** sanskrit recouvrait des réalités hétérogènes : résultat du traitement des sonantes voyelles **m̥* et **n̥*, ancien **e* entraînant la palatalisation d'une vélaire antérieure (***k^{w}*e>* skr. ***ca***), ancien **o* maintenant au contraire l'articulation occlusive de la vélaire antérieure (cf. skr. ***kátara-*** en regard de gr. *pótero-* < k^{w}***otero-***), etc. C'est ainsi un modèle de reconstruction par distinction qui a définitivement prévalu sur ce point précis de la doctrine, conduisant à mettre en évidence, en indo-européen, le statut morpho-phonologique fondamental de la voyelle apophonique **e/o*.

La reconstruction des significations, illustrée avec tant de brio par les travaux d'É. Benveniste, obéit aux deux mêmes cas de figure. Face à deux morphèmes formellement identiques, mais distincts quant à leur sens, elle opère très souvent par synthèse, ramenant les deux formes à

l'unité grâce à une opération de généralisation sémantique. Sur le plan évolutif, cela revient à postuler qu'une entité lexicale ou grammaticale donnée a connu historiquement une situation de polysémie, débouchant ou non ultérieurement sur une scission homonymique[4]. Dans le cas le plus frappant, les formes attestées dans une ou plusieurs langues apparentées sont des contraires ou des contradictoires, tout en présentant une identité formelle : ainsi hittite *dattari* « il est pris » et lat. *datur* « il est donné », dont l'unification est possible à condition de supposer, pour la racine **deh$_3$-* habituellement traduite par « donner », un sens premier tel que « saisir » ou « échanger » (Benveniste, 1951 = 1966, 316). Cette analyse est rendue hautement vraisemblable par tout ce que l'on sait des pratiques du don contraignant dans les sociétés primitives ; voir aussi le sens de fr. *louer* = « donner/prendre en location ». De même, la réduction d'autres polysémies mettant en jeu des inversions actantielles, proposée par F. Bader, 1984 (racines à « diathèse double » type **kel-* « entendre »/« appeler », **leuk-* « voir/briller », etc.), trouve une confirmation dans l'existence bien assurée, dans les langues modernes, de radicaux dits neutres ou symétriques : cf. fr. « il casse la branche »/« la branche casse », « il claque la portière »/« la portière claque », voir aussi les nombreux emplois plus ou moins néologiques tels que « je déprime » face à « cette corvée me déprime », et (à l'inverse) « il démarre la voiture » face à « la voiture démarre ». D'autres réductions, tout aussi spectaculaires du fait qu'elles rapprochent des termes appartenant à des champs sémantiques différents, sont rendues possibles quand une entité lexicale a donné lieu à des emplois métaphoriques. F. Bader (cours à l'ÉPHÉ, 1989-1990) met en lumière un phénomène de ce type à propos d'une racine **seh$_2$-*, d'abord appliquée à la technique du liage (skr. *syáti*, all. *Seile*, etc.), puis utilisée métaphoriquement dans des termes concernant la composition poétique et musicale, notamment à l'initiale de noms propres anciennement motivés comme *Homère* et *Hésiode*.

On signalera qu'il n'est malheureusement pas possible d'accorder une égale vraisemblance à l'ensemble des polysémies qui ont été, parfois abusivement, attribuées à l'indo-européen, notamment par le biais des racines servant d'entrées aux dictionnaires étymologiques (on pourrait ainsi critiquer maints regroupements opérés chez Pokorny, 1959-1969).

---

4. On considère qu'il y a homonymie (fortuite ou historiquement acquise) entre deux signes de la langue quand leur forme est identique et que la signification de l'une ne se laisse pas (ou plus) calculer logiquement à partir de l'autre (ex. : fr. *grève*$_1$ « plage de sable » et *grève*$_2$ « arrêt de travail », issus d'un même étymon, mais clairement ressentis aujourd'hui comme deux mots distincts). En lexicographie, le tri entre cas de polysémie et cas d'homonymie est loin d'être simple même pour une langue moderne et donne lieu à des options divergentes selon les dictionnaires, portés soit à privilégier un point de vue historique, soit un point de vue strictement synchronique (voir Gardes-Tamine 1988, I, 109-110).

Corollaire des phénomènes de spécialisation diachronique, la reconstruction sémantique par unification s'oppose stratégiquement à la reconstruction par distinction, plus difficile à étayer, mais fort intéressante en ce qu'elle contribue à récuser une idée simpliste : celle d'une langue indo-européenne «primitive» au stock lexico-grammatical restreint, qui ne se serait enrichi et diversifié que progressivement, au cours du temps. La reconstruction par distinction permet en effet de montrer qu'une entité lexicale donnée résulte de la fusion de deux radicaux initialement séparés, phénomène bien attesté dans les traditions linguistiques (cf. le cas du français qui tend à confondre *décrépit* <lat. *dēcrepitus* «usé par l'âge» et *décrépi* «dont on a enlevé le crépi» : Saussure, 1916-1972, 119-120)[5]. Ainsi, Benveniste a-t-il par exemple postulé que gr. *dómos* «maison» résulte de la confluence tardive d'un dérivé thématisé d'un nom-racine **dom*, terme institutionnel désignant une institution politique, un cercle de parenté, avec un dérivé thématique de la racine **dem-*(h_2)- «construire» (1955, 15-29 ; l'article contient des explications analogues pour d'autres racines). Pour J.-C. Milner, la démonstration benvenistienne par excellence se signale précisément par la mise en évidence qu'un objet linguistique attesté comme unité doit en réalité être compté pour deux (Milner, 1985, 319) : affirmation peut-être un peu rapide, étant donné les nombreux cas d'étymologie par unification qui parsèment le *Vocabulaire des institutions indo-européennes* (l'exemple de **deh*$_3$- mentionné *supra*, peut paraître tout aussi typique de la méthode reconstructive de Benveniste[6]). Ce qui est incontestable, c'est que celui-ci a exprimé sans ambiguïté le souci de définir «les critères qui permettent d'identifier l'une à l'autre des significations non pareilles» (1955, 14), en vue de fonder et d'équilibrer, en sémantique reconstructive, les démarches d'unification et de distinction.

---

5. Citons également *paon* et *pavaner*, spontanément rapprochés aujourd'hui dans la conscience des francophones alors que le premier est issu de lat. *pavo*, *pavonis*, et que le second est dérivé de *pavane*, lui-même emprunté au XVIe siècle à l'italien *padana* «de Padoue». Le langage enfantin peut présenter plus ou moins fugitivement des phénomènes du même genre : plus d'un enfant de 3 à 5 ans utilise, pour un temps, un signifiant neutralisant /KaTo/ portant manifestement à la fois les sens de «gâteau» et de «cadeau».

6. Exemple qu'on pourrait enrichir de beaucoup d'autres cas de réduction de polysémie «en vue de restaurer l'unité fondamentale de la signification», voir notamment le cas de la racine **med-*, qui présente des sens attestés très divers dans lat. *medeo* (*medeor*) «je guéris», osque *meddíss* «iudex», irl. *midiur* «je juge», gr. *médomai* «je prends soin de», lat. *meditor* «je médite, je réfléchis, je m'exerce», etc., et qui donne lieu de la part de Benveniste à une réflexion méthodologique sur ce mode de reconstruction sémantique (1969, II 124-130).

2.3. Statut des racines et des autres éléments reconstruits

Deux positions principales s'affrontent quant au statut des entités reconstruites. Pour les uns, partisans de l'école réaliste ou substantialiste, la méthode aboutit à la restitution de formes « phonétiques » complètes et concrètes, voire de paradigmes entiers, assignables à peu de réserves près à l'idiome d'origine (voir une profession de foi de ce genre, qu'on pourrait qualifier de post-schleichérienne, chez Szemerényi, 1980, 29). Pour d'autres, tenants de l'école formaliste ou algébriste, une forme reconstruite est au contraire à considérer comme une synthèse d'affirmations abstraites portant sur des parties de mots, comme une formule toujours révisable résumant un faisceau de conjectures, formule dont rien ne permet d'affirmer avec certitude qu'elle a été un jour pourvue, telle quelle, d'existence linguistique (par exemple Saussure, 1916 = 1972, 301).

Contrairement à ce qu'on pourrait croire à première vue, l'option algébriste, permettant le calcul et la prédiction, moins naïvement attachée aux phénomènes de surface, a permis les performances heuristiques les plus notoires et les spéculations les plus rentables. Illustrée par le jeune Saussure dans son *Mémoire* paru en 1878, elle a procuré à la grammaire comparée l'une de ses découvertes les plus marquantes, dont la communauté scientifique de l'époque a néanmoins eu beaucoup de peine à reconnaître le bien-fondé. Sur la base d'un raisonnement analogique dont il sera question plus bas, Saussure démontrait prophétiquement l'existence, en indo-européen, de deux phonèmes supplémentaires notés *A* et *O* : il se souciait essentiellement de spécifier leur fonctionnement dans les alternances grammaticales, mettant délibérément entre parenthèses le problème de leur contenu phonétique. Longtemps contestée, la réalité des « coefficients sonantiques » (plus tard appelés « schwa » et enfin « laryngales ») fut prouvée de manière externe, près de cinquante ans plus tard et bien après la mort de Saussure, par le témoignage du hittite nouvellement déchiffré.

Un tel exemple est rare, et trop rares aussi les occasions où de nouveaux matériaux permettent d'évaluer les doctrines en concurrence. De fait, la rigueur progressivement acquise de la méthode comparative ne permet pas d'oublier que l'objet qu'elle s'est assigné est dépourvu d'existence référentielle, l'indo-européen en tant que tel ne nous ayant pas été livré par une tradition écrite. Alors que beaucoup de discours scientifiques décrivent (ou prétendent décrire) des états de choses ou des états de fait, la grammaire comparée a ceci d'original que *par définition*, elle (re)crée un objet, la langue-source, qui n'est accessible que par l'acte de reconstruction lui-même, aucune « évidence perceptive » n'existant préalablement au travail du savant. Dès lors, les notions fondamentales de « système phonologique », de « racine » ou de « cellule morphologique » indo-européens se trouvent au confluent d'une série de problèmes relevant de l'épistémologie autant que

de la linguistique. Reconstruire un phonème, un suffixe, ou une racine (au sens d'élément indécomposable ultime dégagé par l'analyse morphologique, et véhiculant le sens lexical d'un mot quelconque), c'est bien sûr postuler une forme dont on suppose qu'elle a eu une quelconque réalité dans une synchronie antérieure. Mais c'est, tout autant, faire un acte (créatif) de modélisation, visant à maîtriser la diversité du matériau linguistique observable [7]. C'est aussi et surtout forger, sous la forme d'un sténogramme synthétisant une série d'hypothèses, un outil conceptuel efficace, au service de tâches bien précises. La démarche reconstructive a ainsi permis, à une date relativement récente, le déchiffrement de plusieurs idiomes reconnus comme étant de souche indo-européenne (hittite, grec mycénien, tokharien) ; elle a servi d'auxiliaire à la description de l'histoire des mots, et a autorisé la formulation d'hypothèses culturelles sur l'habitat, l'organisation sociale et économique, la mythologie et les croyances religieuses d'une civilisation préhistorique. Quel que soit le degré de « réalisme » des formes reconstruites, elles sont un accessoire irremplaçable des sciences historiques, comme en témoignent à leur manière aussi bien les travaux des mycénologues que ceux de Georges Dumézil.

2.4. Langue reconstruite et « langue » au sens saussurien

L'activité de modélisation qui est à l'œuvre en grammaire comparée, donnant accès, par les moyens indirects qu'on vient de voir, à l'idiome ancestral perdu, est à la source du concept de « langue » tel qu'il a été généralisé en linguistique du XX[e] siècle à partir du *Cours de linguistique générale*. Indépendamment de l'opposition nécessaire entre points de vue synchronique et diachronique, la filiation des deux notions se situe au plan cognitif : elle est relative à la saisie des données linguistiques par l'intellect, à la perception des identités et des différences. Car les notions de langue reconstruite comme de langue tout court s'élaborent grâce aux mêmes procédures de sélection et d'abstraction. Toutes deux reflètent une double exigence : d'une part, l'appréhension unificatrice des variantes linguistiques observées (variations interlangues dans le cas de l'indo-européen, allomorphies – libres ou conditionnées, dialectales ou idiolectales – internes à un système linguistique dans le cas du concept de langue) ; d'autre part et corollairement, le tri ou la distinction des signes homophones sur la base d'un critère stable, le critère sémiotique. Elles se heurtent par

7. Il convient d'évaluer sous cet angle-là, à mon sens, la théorie benvenistienne de la racine indo-européenne (Benveniste, 1935), dont les adversaires ne voient pas que le but est de fournir une *définition opératoire* du fonctionnement radical, plutôt que de pointer des entités linguistiques concrètes (Reichler-Béguelin, 1986).

ailleurs aux mêmes problèmes de validation, soupçonnées périodiquement l'une et l'autre d'être des artefacts de grammairiens au lieu de désigner une « réalité » langagière soit révolue, soit contemporaine.

3. Obstacles et limites

Même si beaucoup de comparatistes, réalistes aussi bien que formalistes, affichent une grande confiance dans la méthode qu'ils appliquent, et accordent en conséquence un haut degré de certitude ou de plausibilité aux éléments restitués de la protolangue, plusieurs facteurs sont de nature à tempérer cet optimisme, en même temps qu'ils invitent à méditer, plus généralement, la façon dont le savoir humain construit ses objets.

Certains de ces facteurs relèvent des limites inhérentes à la méthode elle-même (**3.1.**) ; d'autres, aux caractéristiques mêmes de l'objet décrit (**3.2.**) ; d'autres enfin tiennent à certaines idées reçues des chercheurs, qui sont parfois portés à injecter plus ou moins consciemment dans l'idiome ancestral soit certains préjugés sur la langue primitive, soit certains idéaux de simplicité ou de logique étrangers au fonctionnement des langues naturelles (**3.3.**).

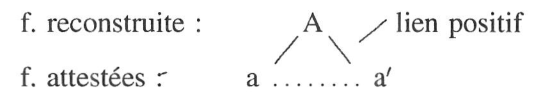
3.1. Place(s) de la subjectivité du chercheur dans l'exercice de la méthode

On l'a déjà esquissé plus haut (cf. **2.2.2.1.**), comparaison et reconstruction apparaissent chez beaucoup de comparatistes comme les deux phases d'une même opération, et reçoivent leur sens l'une de l'autre :

> Si le seul moyen de reconstruire est de comparer, réciproquement la comparaison n'a pas d'autre but que d'être une reconstruction. (Saussure, 1916, 299)

Pourtant, cela ne doit pas masquer le statut bien distinct des relations qui sont envisagées entre les termes comparés d'une part, et entre termes comparés et terme reconstruit d'autre part. Saussure tenait à souligner qu'il existe entre forme reconstruite et formes attestées un lien « positif, historique », alors qu'entre les formes comparées, la relation « horizontale » est « simplement établie par l'esprit, ineffective » (Reichler-Béguelin, 1980,

24) ; cette situation épistémologique bien particulière, qui consiste à affirmer, sur la base d'une relation établie subjectivement, l'existence d'une relation positive, peut être symbolisée par le schéma suivant :



Une position différente, privilégiant non le statut ontologique de l'objet-langue, mais le point de vue rétrospectif de l'observateur confronté, d'une part, à la positivité des données linguistiques attestées, et d'autre part, au statut évanescent de l'indo-européen, s'exprime en revanche dans les travaux d'un autre grand comparatiste, A. Meillet. Ce dernier a manifesté beaucoup de réserve quant à la possibilité, voire quant à la nécessité, d'atteindre la langue reconstruite, qui n'est pour lui qu'un moyen commode de se représenter l'histoire des langues, et non un but en soi (pour plus de détails, voir Reichler-Béguelin, 1988) :

> L'indo-européen est inconnu, et les concordances *sont la seule réalité* qu'ait à étudier le comparatiste. La grammaire comparée *n'a pas pour but de reconstruire l'indo-européen.* (1937, VIII, je souligne)

Il n'est sans doute pas utile de vouloir trancher à tout prix entre ces deux variantes du formalisme, dont l'une privilégie épistémologiquement les reconstructions, l'autre les comparaisons. On retiendra simplement ici que formes comparées et formes reconstruites peuvent, selon les cas, passer pour le lieu d'une réalité, d'une positivité, ou au contraire être ressenties comme lieux d'intervention possibles de la subjectivité du chercheur : tels sont en définitive les paradoxes internes de la grammaire comparée, les limites spécifiques à la démarche elle-même et qui invitent à ne pas prétendre « tirer de la méthode plus qu'elle ne peut donner » (Hjelmslev, 1966, 166).

3.2. Limites découlant des données elles-mêmes

D'autres obstacles, mis en évidence de manière répétée par les comparatistes eux-mêmes (voir, par exemple Meillet, 1925), sur lesquels je passerai donc rapidement, résultent de l'état fragmentaire et imparfait de notre documentation. D'abord, les langues de souche indo-européenne ne sont pas toutes attestées, certaines ayant disparu sans laisser de traces ; d'autres comme le vénète, le gaulois, le celtibère, ne nous sont parvenues que sous forme de bribes. Des langues classiques elles-mêmes, on ne connaît que la partie émergée de l'iceberg : essentiellement une langue

littéraire plus ou moins normalisée, à laquelle, de surcroît, le médium graphique fait partiellement écran. Les écarts chronologiques dans les témoignages sont frappants : alors que le grec mycénien et le hittite sont connus avant le 1[er] millénaire av. J.C., les premiers documents lituaniens ne datent que du XVI[e] siècle de notre ère, sans que l'on sache clairement mesurer les incidences de cette disparité sur les résultats de la grammaire comparée. Enfin, l'exemple de la famille romane, dont on a la chance de posséder la langue d'origine, nous enseigne que les idiomes issus d'une même tradition peuvent connaître des innovations et des pertes parallèles (ainsi les langues romanes ont-elles innové par rapport au latin en acquérant toutes la catégorie de l'article). Il en découle un problème insoluble pour la méthode comparative, portée par définition à considérer comme pan-indo-européennes *l'ensemble* des caractéristiques partagées par les langues-filles, et, symétriquement, privée à tout jamais des éléments qui auraient été éliminés sans exception par celles-ci. La prudence s'impose notamment quand deux formes comparables appartiennent toutes deux, dans leurs langues respectives, à des formations productives : elle sont alors susceptibles d'être l'une comme l'autre des créations analogiques relativement récentes. Il résulte de tout cela qu'il n'est possible de restituer de l'indo-européen que des fragments approximatifs, d'époques diverses, bien éloignés de la totalité systématique et synchronique que nous avons pris l'habitude d'appeler « langue ».

D'autres difficultés, moins souvent signalées, tiennent à la structuration même du langage humain et à l'interdépendance des différents niveaux d'analyse linguistique. Ainsi, en dépit du fait que les évolutions phonétiques sont *a priori* plus propices à l'observation objective que les changements sémantiques, parce que moins directement liées à des facteurs culturels, la reconstruction d'un système phonologique primitif n'a pas conduit à une solution unique et définitive : ni le nombre, ni la nature phonétique exacte des phonèmes indo-européens ne peuvent être fixés de manière indubitable. Dans les cas douteux, il est frappant de constater que la décision de reconstruire ou non un phonème dépend étroitement des rapprochements morphologiques que l'on est prêt ou non à opérer grâce à lui. Si l'on prend l'exemple célèbre et longtemps controversé des laryngales (cf. **2.3.**), leur reconstruction par Saussure n'a été possible que grâce à l'hypothèse, purement *morphologique*, d'une structuration régulière des racines indo-européennes, les racines à voyelle longue de forme *stā-*, *dō-*, étant interprétées comme issues de **steX-*, **deY-* (= $*steh_2$-, $*deh_3$-), par assimilation à un schéma canonique C_1eC_2 bien représenté par ailleurs : **ten-*, **dem-*, $*b^her$-, etc. De manière analogue, la décision de poser d'autres laryngales que le trio plus ou moins incontesté h_1, h_2, h_3, dépend entièrement de la postulation d'autres régularités, plus ponctuelles cette fois, mais relevant toujours de la « première articulation du langage » au

sens de Martinet. Acceptera-t-on, par exemple, de restituer *E^y, laryngale palatale, et *A^w, laryngale labio-vélaire (admises par Haudry, 1984, 17) ? Ce sera pour rendre compte de certains doublets où alternent une voyelle longue et une séquence voyelle brève + semi-voyelle, ainsi dans **d^hē-/*d^hey-* « allaiter, téter » et **dō-/*dow-* « donner, recevoir », et la saisie de ces doublets comme unités morphologiques anciennes, vrai lieu de la controverse, conditionne l'hypothèse phonologique. Voilà pour le *nombre* des unités.

Quant à leur contenu phonétique, il faut noter le recours récent aux parallèles typologiques comme moyens de contrôle, sur la base des travaux de Gamkrelidze et Ivanov (voir la synthèse détaillée et la critique de Mayrhofer *in* Cowgill et Mayrhofer, 1986, 92-98). La comparaison typologique a en effet conduit à une révision du système des occlusives postulé pour la protolangue indo-européenne : l'idée classique d'un système à séries sonores-sourdes-aspirées est remplacée par celle d'un système à séries glottalisées-sonores-sourdes (ces deux dernières avec allophones aspirés), apte à justifier l'absence, ou du moins la grande rareté, d'un phonème **b* indo-européen. Relevons que cette conception, actuellement discutée, ne met pas en péril l'acquis étymologique, et notamment ce qu'on sait de la morphologie ; elle conduit toutefois à considérer comme conservatrices, en matière de traitement des occlusives, les langues dites « à mutation consonantique » comme le germanique, l'arménien et le hittite.

L'imbrication des conjectures, qui rend pour ainsi dire impossible une unification du savoir en grammaire comparée, se fait sentir davantage encore lorsqu'on passe à l'évaluation des rapprochements morpho-sémantiques eux-mêmes. Si l'on entend par exemple faire remonter à une même racine **b^her-* « porter » deux formes comme lat. *ferō* « je porte » et *fors* « le hasard », il faut faire une supposition d'ordre sémantique et admettre que le second, ancien nom d'action de forme **b^hr̥-t(i)-*, a connu un changement de sens : « le fait de porter (ou : d'apporter) » > « le hasard ». Refusé par Ernout et Meillet dans leur *Dictionnaire étymologique de la langue latine*, ce rattachement rencontrerait pourtant des faits parallèles en grec (où *sumphorá* est en relation très claire avec *phéro:*) et en germanique (voir le signifié du verbe de même racine v. sax. *giburian* = « arriver, se rencontrer »). Quoi qu'il en soit, dès que des significations sont en jeu, le problème se pose de savoir ce qu'est une « évidence » en grammaire comparée. Opérer ou non une reconstruction morpho-sémantique, cette décision dépend bien souvent de l'idée qu'on a du fonctionnement grammatical de la langue d'origine, des dérivations de sens qu'on y estime plausibles, et aussi du contexte socio-culturel dans lequel on la situe. Ainsi, rapprocher gr. *mántis* de *me:núo:* « je déclare, j'annonce » plutôt que de *maínomai* « je délire », comme l'a suggéré naguère M. Casevitz, implique non seulement un raisonnement grammatical, mais aussi et surtout un enjeu culturel : il

s'agit de renoncer, pour un état de langue et de civilisation donné, à la conception romantique du devin vaticinant dans la transe, de remplacer le stéréotype du devin visionnaire par l'idée plus neutre, plus rationnelle, d'un personnage dont le rôle serait d'annoncer, de dévoiler, d'avertir (cf. le sens des composés *kakómantis*, *oio:nómantis*, *oneirómantis*, etc.). Une telle idée prête bien entendu à discussion[8].

3.3. Obstacles liés aux préjugés éventuels des chercheurs

3.3.1. Métaphores du changement

Pour faire de l'étymologie une pratique rigoureuse, pour élaborer une image scientifiquement fiable de l'évolution linguistique, il est bien connu que les premiers comparatistes se sont appuyés sur le modèle des sciences de la nature[9]. Or, plus d'une fois, en même temps qu'elle était facteur de progrès, la métaphore inspiratrice induisit une vision partiellement erronée du phénomène langagier. Ainsi en est-il du modèle évolutionniste de Schleicher, qui, botaniste de formation et fortement influencé par les travaux de Linné, de Cuvier et de Darwin, contribua décisivement à la mise au point de la méthode reconstructive dans ses travaux de détail, tout en concevant par ailleurs une théorie de l'évolution linguistique curieusement partiale et déconnectée des locuteurs. Selon Schleicher en effet, les langues, assimilées à des organismes naturels, connaissent, en dehors de toute influence humaine, croissance, apogée et déclin : chacune passe tour à tour, dans une phase de perfectionnement, par les types isolant, agglutinant, puis flexionnel, avant de se dégrader au cours d'une nouvelle phase dite de «dégénérescence». Cette théorie, tout en conférant une prééminence injustifiable au type flexionnel dont relève l'indo-européen, avait pour particularité de mettre totalement entre parenthèses le caractère social du langage, dont l'étude devenait dès lors une branche des sciences naturelles plutôt que des sciences humaines. Le nom de Schleicher reste, d'autre part, attaché à la *Stammbaumtheorie* (théorie de l'arbre généalogique), qui

8. Voir Casevitz, 1992. Même quand l'enquête est possible, et même quand le rattachement d'un lexème à tel ou tel radical ne fait pas de doute, les dérivations sémantiques sont loin d'être toujours transparentes. Citons un exemple tout simple, mais révélateur : dans certaines régions de Suisse romande, on désigne un café au lait du terme de *renversé*, sans que les locuteurs sachent au juste s'il faut attribuer cette appellation à une particularité technique (le café s'écoulant du percolateur serait reversé d'une petite tasse dans une plus grande) ou à une inversion de proportions des composantes du «café crème» standard (au contraire de ce dernier, le *renversé* contient beaucoup de lait, peu de café).

9. Pour plus de détails, voir les travaux cités note 2.

groupait de manière commode, mais excessivement rigide, les dialectes indo-européens en branches et rameaux nettement scindés (d'où les notions, ultérieurement contestées, de « langues communes » balto-slave ou italo-celtique) ; après avoir dominé longtemps, cette doctrine dut céder le pas, dans la seconde moitié du XIX[e] siècle, devant les progrès de la dialectologie et de la géographie linguistique, qui vinrent confirmer la théorie concurrente préconisée par J. Schmidt (*Wellentheorie*). Selon cette dernière, la propagation de chaque changement linguistique s'opère de manière autonome, à la manière d'ondes s'éloignant concentriquement à partir de points d'origine distincts. D'où une vision beaucoup plus complexe des processus de scission dialectale, une vision plus nuancée également de la langue-source indo-européenne, elle-même non homogène et « toujours déjà » dialectalement différenciée. De son côté enfin, le modèle « mécaniste » des néo-grammairiens (Brugmann, Osthoff, Leskien, Paul, etc.) fit triompher la méthode positiviste en linguistique historique, grâce à la démonstration de l'absolue régularité des changements phonétiques, dont l'action ne peut être contrebalancée que par l'analogie. Il eut néanmoins l'inconvénient de favoriser une vue atomiste des faits de langue, conçus de manière exclusivement historiciste, et indépendamment des micro-systèmes qui les contiennent. Tout se passe donc comme si chaque métaphore du changement ne permettait d'éclairer qu'un des aspects du phénomène linguistique, en omettant de rendre justice aux autres.

3.3.2. De quelques idées reçues à propos de la langue primitive

3.3.2.1. Répugnance pour l'abstraction

D'autres exemples, plus ponctuels, montreront à quel point les conceptions générales des linguistes sont susceptibles d'influencer leur manière de reconstruire l'indo-européen. Un préjugé de type « primitiviste » a ainsi conduit Specht, après quelques autres, à soutenir que la langue d'origine ne pouvait exprimer des concepts abstraits autrement que par le biais de « personnifications » :

> « *Bereits Usener hat in seinen Götternamen 364 ff. die Anschauung vertreten, dass die Abstrakta jüngere Gebilde sind und sich aus Personifikationen entwickelt haben. Im Anschluss daran gebührt Kretschmer das Verdienst, Glo. 13, 101 ff. gezeigt zu haben, dass die ursprüngliche religiöse Anschauung der Indogermanen in einer Art von Animismus wurzelte. [...]* » « *Trotzdem kann, im ganzen gesehen, kein Zweifel darüber bestehen, dass die Abstrakta erst aus belebt gedachten Begriffen hervorgegangen sind.* » (1944, 386 et 391)

Pétri de conceptions animistes, l'indo-européen n'aurait donc, d'après cet auteur, connu que des « concrets », et tous les noms d'action récents

devraient être interprétés comme d'anciens noms d'agent. Cette doctrine, on le devine, n'a guère eu de retentissement; elle est d'autant plus contestable qu'elle méconnaît entièrement la fonction linguistique des abstraits, qui ne servent pas à désigner des notions intangibles, mais bien à assurer un certain type de translations syntaxiques (adjectif → substantif dans le couple lat. *innocens/innocentia*, verbe → substantif dans *comparāre/comparātiō*). Toute langue fait usage de telles translations, aucune ne saurait donc être constituée uniquement par un stock de signes à valeur «concrète», comme le voulait Specht dans une vision purement dénotationnelle ou référentialiste du langage.

3.3.2.2. Répugnance pour la polysémie

Le linguiste en général, et le comparatiste en particulier, est parfois victime à son insu d'illusions qui risquent de lui masquer la réalité des mécanismes langagiers. À l'heure actuelle, seule une observation vigilante du fonctionnement des langues vivantes peut éviter aux indo-européanistes certaines erreurs d'appréciation, comme celle consistant à méconnaître, dans les procédures de reconstruction de l'indo-européen, les phénomènes de polysémie, de variation, de neutralisation d'oppositions qui caractérisent l'ensemble des langues naturelles. Par exemple, dans un ouvrage subtil et documenté étudiant la famille lexicale de lat. *cornu*; gr. *kára*: «tête» et *kéras* «corne»; skr. *śíraḥ* «tête» et *śŕ̥ṅga-* «corne», etc., A. J. Nussbaum (1986) commence par remettre en cause l'idée reçue selon laquelle deux formes indo-européennes de base **ker-* et **kerh$_2$-*, l'une aniṭ l'autre seṭ, auraient toutes deux comporté une valeur sémantique plus ou moins indifférenciée «tête»/«corne». Il n'admet pas non plus que l'une et l'autre signification aient pu résulter de la spécialisation d'un signifié reconstruit plus général «extrémité supérieure du corps». Or, de manière générale, étant donné ce qu'on sait des polysémies tropologiques, il reste problématique que l'auteur, au long de son ouvrage, exclue *a priori*, pour les items lexicaux considérés, la possibilité de changements de sens à caractère métonymique, non liés à des phénomènes de dérivation.

Ce qui est contestable ici, ce n'est pas tant le résultat des analyses proposées que le type d'argumentation qui les fonde. Grande est en effet la tentation soit d'imaginer une langue-source à lexique pauvre et maximalement polysémique (cf. *supra* **2.2.2.2.**), soit au contraire d'injecter dans l'idiome originel un idéal de monosémie et de proscription des variantes de signifiant qui n'est pleinement réalisé que dans les langages formels, en logique ou en mathématique : C. Bally parle ainsi du «double idéal que poursuit la communication : la *linéarité* et la *monosémie*», idéal qui, il faut le redire après Bally, n'est pourtant «jamais atteint par aucune langue» (1935, 141).

3.3.2.3. L'ordre (chrono)logique des constituants de la phrase

En syntaxe également, certaines idées reçues peuvent conduire à fausser l'image du prototype par excès d'idéalisation. Ainsi l'hypothèse que dans la phrase simple indo-européenne, «... l'ordre des éléments de l'énoncé est *chronologique* : l'origine, puis les circonstances puis le procès, avec son terme; enfin, le but ou la conséquence de ce procès» (Haudry, 1984, 107). Ne retombe-t-on pas ici dans le cliché naturalisant, un peu désuet, d'une syntaxe à fondement «(chrono)logique», reflétant l'ordre du monde plutôt que les contraintes pragmatiques de l'interaction verbale, c'est-à-dire les nécessités de mise en relief informationnelle et de mémorisation ?

On constate aussi par ailleurs, à travers de tels exemples, qu'un des grands problèmes qui se pose au linguiste est de ne pas confondre «simplicité» de l'explication qu'il propose, et principe unique et immuable de l'économie langagière.

4. Conclusion

Quels que soient par ailleurs les mérites des travaux dont ils sont extraits, les cas mentionnés plus haut peuvent inspirer quelques réflexions méthodologiques, qui serviront de conclusion à cette sommaire étude. En linguistique historique comme en linguistique générale, le plus difficile est de ne pas imposer d'avance un modèle normatif à la langue qu'on cherche à connaître, étant entendu que dans n'importe quelle description, y compris celle de l'indo-européen, «les représentations intériorisées» qu'on a de l'objet jouent un rôle essentiel (Apothéloz, 1983, 3). C'est pourquoi les comparatistes, toujours en lutte avec les tentations cratylienne ou logiciste, ont à s'interroger en permanence sur les moyens d'accréditation des reconstructions qu'ils opèrent, sur les moyens susceptibles de réduire, dans leur discours, la part de l'argument d'autorité, ainsi que celui, plus ou moins incontrôlable s'il n'est pas étayé, de la «vraisemblance». En privilégiant l'argumentation proprement linguistique, fondée sur une prise de conscience claire et exhaustive du fonctionnement des langues naturelles, en contribuant à l'étude des phénomènes variationnels, et, selon le vœu de Meillet, à celle du «comportement des sujets parlants envers les règles», ils peuvent mettre leur expérience irremplaçable au service d'une meilleure connaissance du changement linguistique, qui conserve beaucoup de ses mystères.

Bibliographie

APOTHÉLOZ D. (1983). – « Éléments pour une logique de la description et du raisonnement spatial », *Degrés* 35-36, 19 p.

AUROUX S. (éd.) (1990-). – *Histoire des idées linguistiques*, 3 tomes, Bruxelles, Mardaga.

BADER F. (1984). – « Autour de Polyphème le Cyclope à l'œil brillant : diathèse et vision », *Die Sprache* 30/2, p. 109-137.

BALLY C. (1935). – *Le Langage et la Vie*, Zurich.

BENVENISTE E. (1935). – *Origines de la formation des noms en indo-européen*, Paris.

– (1951). – « Don et échange dans le vocabulaire indo-européen », *L'Année sociologique*, 3e série, t. II.

– (1955). – « Homophonies radicales en indo-européen », *BSL* 51, p. 14-41.

– (1966). – *Problèmes de linguistique générale*. I, Paris.

– (1969). – *Le Vocabulaire des institutions indo-européennes* (2 tomes), Paris.

BONNET C. et TAMINE J. (1985). – *Quand l'enfant parle du langage*, Bruxelles, Mardaga.

CASEVITZ M. (1992). – « *Mantis* : le vrai sens », *REG* CV, p. 1-18.

COWGILL W. et MAYRHOFER M. (1988). – *Indogermanische Grammatik I, Einleitung, Lautlehre*, Heidelberg.

GARDES-TAMINE J. (1988). – *La Grammaire* (2 tomes), Paris.

HAUDRY J. (1984). – *L'Indo-Européen*², Paris.

HJELMSLEV L. (1966). – *Le Langage*, Paris.

LEROY M. (1971). – *Les Grands Courants de la linguistique moderne*², Éditions de l'Université de Bruxelles.

MEILLET A. (1925). – *La Méthode comparative en linguistique historique*, Oslo et Paris.

– (1937). – *Introduction à l'étude comparative des langues indo-européennes*⁸, Paris.

MILNER J.-C. (1985). – « Sens opposés et noms indiscernables : K. Abel comme refoulé d'É. Benveniste », *in* Auroux (Sylvain) *et al.* (éds), *La Linguistique fantastique*, Paris, p. 311-323.

MOUNIN G. (1974). – *Histoire de la linguistique des origines au XXe siècle*³, Paris.

NUSSBAUM A. (1986). – *Head and Horn in Indo-European*, Berlin et New York, Walter de Gruyter.

OLENDER M. (1989). – *Les Langues du paradis. Aryens et Sémites : un couple providentiel*, Paris.

POKORNY J. (1959). – *Indogermanisches etymologisches Wörterbuch*, Berne.

REICHLER-BÉGUELIN M.-J. (1980). – « Le consonantisme grec et latin selon F. de Saussure : le cours de phonétique professé en 1909-1910 », *CFS* 34, p. 17-97.

– (1986). – « Description et représentation dans le discours de la linguistique historique : l'exemple des racines indo-européennes », *Actes du Colloque sur la Description*, Fribourg, juin 1986, à paraître.

– (1988). – « La méthode comparative de Meillet : statut et légitimité des reconstructions », in *Antoine Meillet et la linguistique de son temps*, S. Auroux, éd., *Histoire Epistémologie Langage* X/2, p. 11-24.

– (1989). – « La notion de preuve en grammaire comparée des langues indo-européennes », *in* Reichler-Béguelin M.-J. (éd.), *Perspectives méthodologiques et épistémologiques dans les sciences du langage*, Actes du Colloque de Fribourg, 11 et 12 mars 1988, Berne, Peter Lang, p. 115-138.

ROUSSEAU J. (1984). – « La racine arabe et son traitement par les grammairiens européens (1505-1831) », *BSL* 79, p. 285-321.

SAUSSURE F. de (1878). – *Mémoire sur le système primitif des voyelles dans les langues indo-européennes*, Leipzig (cité d'après le *Recueil des publications scientifiques de F. de Saussure*, Genève-Paris 1984 = Genève 1922).

– (1916). – *Cours de linguistique générale*, éd. par C. Bally et A. Sechehaye, Paris, Payot (éd. critique par T. De Mauro, 1972 ; cité aussi dans l'édition critique et synoptique de R. Engler, Wiesbaden, dès 1967).

SPECHT F. (1944). – *Der Ursprung der indogermanischen Deklination*, Göttingen.

SZEMERÉNYI O. (1980). – *Einführung in die vergleichende Sprachwissenschaft*², Darmstadt.

CHAPITRE 3

LES NOMS DES ARYENS : ETHNIQUES ET EXPANSION

Françoise BADER

1. ENSEMBLE ONOMASTIQUE STRUCTURÉ

Les hommes de langue indo-européenne se sont donné des noms qui forment un ensemble structuré, fragmenté selon un procédé dont leur onomastique offre bien d'autres témoignages : ainsi les anthroponymes mythiques, comme ceux qui structurent chez Homère l'idéologie du κλέον ἄφθιτον définie par Achille, *Iliade* 9, 411-6 : Nestor « qui survit (en sa vieillesse) » ; Achille « qui endeuille le peuple en armes »[* Ἀχί-λᾱος] ; Héraklès « qui obtient sa gloire, son κλέος d'Héra » (et non du combat) ; les toponymes qui comme Delphes/Pythô peuvent avoir un nom qui se rapporte à une particularité physique (ici, le « nombril ») et un autre à un exploit (ici celui d'Apollon contre le serpent Python), ou qui, dans l'espace, peuvent, comme Λήρισαι et Νέον Τεῖχος (Hérodote 1, 149), s'appliquer à deux lieux géographiquement voisins, définis chacun par l'une des deux données de la défense pierreuse, la première naturelle (« le (lieu) des

Pierres»), la seconde culturelle (le mur d'enceinte construit); les formulaires toponymiques dont des exemples ont été donnés dans l'Introduction, § 1.3.

Les ethniques qui, comme par exemple les toponymes retenus dans ces formulaires, sont tous fondés sur un appellatif, sont relatifs à l'expansion territoriale et à l'établissement politique de ceux qui se les sont attribués : «Migrants» (*Pélasges*), «Désirant conquérir» (*Vénètes*), venus d'un «Autre» territoire (*Ārya-*), ils ont fait «Leur» le territoire de conquête (*Suessionēs*, *Suēbī*, *Sabīnī*), pour l'avoir organisé en «réunion, fédération» (*Semnonēs, Sabīnī*), en «totalité politique» (*Teutōnēs*) de «tous les hommes» (*Ala-mannī*, *Slaves*, *Hellènes*), qui transcende l'opposition de l'«autre» et du «soi», des conquérants et des conquis. Leur successivité, telle qu'elle vient d'être résumée, est adaptée de manière iconique au déroulement dans le temps de leur conquête et de leur établissement, et déterminera le plan selon lequel ils seront examinés. L'existence même de cette structure pose un problème sans solution à l'heure actuelle : pour que ces hommes se soient donné l'un de ces noms, il faut qu'ils aient connu l'existence des autres; et il en va de même à une époque moins reculée que celle à laquelle ils renvoient par exemple pour les noms issus du *uēr sacrum* italique (§ 4). Car chacun de ces noms ne reçoit sa valeur sémantique propre que de l'ensemble dont il fait partie. C'est donc cet ensemble, ainsi que l'appellatif qui est à la base de chacun de ces noms, qui fournira la valeur sémantique de ceux-ci.

2. PÉLASGES

Le nom des «*Pélasges*» est fait sur la racine «errer» de πλανάομαι, *$p\c{l}$-h_2-*, qui a eu une forme élargie par *-g-* : *$p\c{l}$ h_2-g-* par exemple dans le nom des «*Planctes*», terres «Errantes» (dans la mer) avant de se fixer, en une antique cosmologie, ou dans le verbe employé pour Ulysse, *Odyssée* 1, 1-2, ὃς μάλα πολλὰ | πλάγχθη, dès le début de l'*Odyssée* (qui peut être lue comme un mythe initiatique d'errance); et *pel-h_2-g-*, dans trois ethniques :

a) Πελάγονες (d'où l'anthroponyme Πελάγων et le toponyme Πελαγονία au Nord de la Macédoine; il est fait, comme Στράβων sur στραβός, sur un *πέλαγος, ου qui est à πέλαγος, ους «mer» (conçue comme «lieu d'errance») ce que par exemple Ἄργος, ου est à Ἄργος, ους, toponyme;

b) le nom des «Pélasges» < *pel-h_2-g- + *-sko-* avec le même suffixe que par exemple *Op-scī* «Osques», et le même phonétisme que μίσγω <**mi-g-+-skō*; autre suffixe en sifflante, *-sti-* (cf. de la racine «être large

et plat» *$p\r{l}h_2$-°sti-* > hitt. *palha-sti-*, gr. παλαστι-αῖος) dans πελαστικόν, nom d'un mur d'Athènes réputé avoir été fondé par les Étrusques, par ailleurs appelés aussi Pélasges (**pelh*$_2$*(g)-* + °*sti-ko-*), et qui a aussi pour nom πελαργικόν, dérivé de :

c) πελαργός qui s'échange avec πελασγός dans la tradition manuscrite, s'agissant des Pélasges d'Attique, et qui est identique au nom des «cigognes». C'est, non un dérivé comme les deux ethniques précédents, mais un composé *πελαγο- +-αργός >* πελαγᾱργός > πελᾱργός, par superposition syllabique, avec -ᾱ- attesté, puisqu'il est condamné, Phryn. 88; il est du type dit dvandva, comme par exemple les anthroponymes Λέπταργος, «Léger et Rapide», Σίμαργος «(au nez) Camus et Brillant», avec un second membre de sens tantôt «rapide» tantôt «brillant», comme de manière générale les termes de la famille de **h*$_2$*er-g-*. De sens «errant [et] rapide» (ainsi que «d'un blanc brillant» pour les cigognes), ce composé peut avoir été simultanément employé comme ethnique et comme nom d'oiseau, à moins que le premier ait été tiré du second. En tout cas, l'oiseau a un symbolisme migratoire clair; et il n'est pas fortuit que les Pélasges, les oiseaux et une forme de πλανάομαι «errer» soient associés dans des textes, *e.g.* Strabon 5, 2, 4 (221); Serv., *ad Aen.* 8, 600, Isid., *Etym.* 9, 2, 74, deux auteurs qui soulignent l'arrivée des Pélasges (d'Italie) par «voiles», *uēlis* (ce qui s'accorde avec le rapprochement étymologique de πέλαργος), au printemps (saison du retour des oiseaux, et du *uēr sacrum*).

Propre au grec, et employé pour les Étrusques dans des conditions qu'il ne nous appartient pas d'examiner, le nom des Pélasges, s'il n'est pas par là même un bon exemple d'un ethnique accessible à la comparaison, montre dans quelles conditions il est possible d'attribuer une étymologie et un sens à un nom propre, en principe dépourvu en tant que tel de l'une et de l'autre, que seul un contexte permet de fonder. Le contexte, dans le cas des semi-mythiques Pélasges, est fourni, comme pour un grand nombre de noms mythiques, par le mythe même, ici celui des oiseaux migrateurs au nom apparenté. Aucun argument de ce genre ne pourra intervenir pour l'explication des autres noms, historiques, de l'ensemble sémantiquement structuré ici proposé; mais un autre argument est commun à l'interprétation du nom des Pélasges et de ceux-ci : le fait que tous ces ethniques sont fondés, dans la langue, sur un appellatif qui en fournit le sens et l'étymologie. Si l'appellatif *πελαγός est masqué par la dérivation et la composition dans le nom grec des Pélasges, ce n'est pas le cas des autres noms ici examinés, celui des Vénètes, fait comme le nom des Pélasges sur une racine verbale, ceux des types Aryens, Sabins, Samnites, Teutons, issus de lexicalisations pronominales.

3. VÉNÈTES

Le nom des *Vénètes* est un dérivé de **h_1w-en-/ *h_1w-n-* (> **wn̥h_1-* par métathèse) qui est apparenté par son sens à d'autres racines de sens « désirer » (comme **h_1w-el->* lat. *uelle*, gr. (ἐ)έλδομαι), « espérer » (gr. (ἐ)έλπομαι), et, signifiant proprement « désirer obtenir d'un autre par des actes », en couple avec **pr-ek-* « désirer obtenir d'un autre par des paroles ».

3.1. Racine et problèmes sémantiques

De la complication sémantique des formes de cette famille, qui tient à ce que le procès qu'indique la racine implique un long déroulement dans le temps, avec des modalités et des finalités diverses, on se contentera ici d'abord de donner trois illustrations :

a) les sens pris par le présent **$h_1$wn̥ -sko-* (**wn̥ h_1-sko-*) : skr. *vā́ñchati* « désirer », comme all. *wünschen* ; mais hitt. *uwanšik-* « futuere », et tokh.A *winās-* / B *winā́sk-* « vénérer » ;

b) les sens pris par le dérivé en **-i-* (élargi parfois en **-i-n-*, **-i-o-*, **-i-ā-*) : got. *wēns* « espoir » ; tokh. A *wañi-* « plaisir » ; skr. *vaní-* « désir, demande » nom d'action et nom d'agent en second membre de composé *upamāti-váni-* « qui agrée un discours », et simple en *-i-n- vanín-* « désirant ; distribuant » (exemple de diathèse factitive de cette racine d'échange) ; av. *vanya-* « victoire » et *-vanya-* « qui vainc » ; gaul. (anthroponyme) *Venitouta*, quelque chose comme « famille et patrie (totalité politique) », avec des formes apparentées aux noms des Vénètes et des Teutons ; v. irl. *fine* « famille » < **$h_1$weniā*, comme lat. *uenia*, qui a un autre sens « grâces accordées » (par les dieux ; et par les hommes : « pardon ») ; v. irl. *fuin* « pardon » < **$h_1$woniā* ;

c) la diversité des sens de l'all. *wünschen*, *gewinnen* « gagner, remporter », *wohnen* « habiter », *gewöhnen* « être habitué à », *Wahn* « espoir, attente » ; *Wonne* « plaisir (sexuel) » ; et cf. v.h.a. *wini* « ami », *winnan* « combattre, s'épuiser de travail », got. *winnan* « souffrir », v.angl. *winn* « travail, peine, combat ».

On cernera ensuite quelques-uns des champs sémantiques de cette racine (qui a donné au français des termes aussi différents que « péché *véniel* » et « maladie *vénérienne* »).

L'un est la vénération, pour laquelle **$h_1$wen-os-* s'applique aux deux partenaires de l'échange, les dieux, skr. *gír-vaṇas-*, *yajñá-vanas-* « qui dési-

rent obtenir des actes de grâces, des sacrifices», et les hommes qui les «vénèrent», lat. *ueneror*, dénominatif en regard de tokh.AB *winās(k)-*; au vocabulaire religieux appartient *$*h_1$won-tā-*, av. *vantā-bərəti-* «acte d'apporter des hommages», vén. *vo.l.te.r.ko.n. vo.n.ta.r.* «par un acte volontaire de vénération», formule allitérante, qui trouve un écho dans la mention homérique du nom des Vénètes (§ 3.3). Un autre champ sémantique est celui du «désir sexuel», qui aboutit à la reproduction, et par elle à la famille : lat. *uenus*, all. *Wonne*, et *Wonne-monat*, nom du mois de mai, qui est le mois de la reproduction; v.irl. *fine* «famille», *coibnes* «allié, membre de la même famille; parenté, consanguinité»; on soulignera ici l'emploi de termes de la famille pour des divinités et des généalogies mythiques : av. *Kavāta-* (**Kavi-vāta-*), nom du fondateur de la dynastie des *Kavi-*; lat. *Venus*, incarnation de *uenus* «acte sexuel», mais aussi mère d'Enée; *Vanes*, nom générique des divinités germaniques de la fécondité. L'objet du désir peut être divers (c'est parfois un «ami» : v.h.a. *wini*, v.isl.*vinr*), et l'obtention comporter bien des modalités, «acte de plaire» (lat. *uenustās*), «argumentation» (v.h.a.*wānida*), et devenir violente quand l'échange est impossible (lat.*uēnor* «chasser»), ou que l'argumentation échoue : elle consiste alors à «conquérir» (all. *gewinnen*), champ sémantique particulièrement bien représenté en avestique : *vanana-* «vainqueur», *vananā-* «victoire»; *haθra-vanya-* «qui vainc d'un coup»; *vīspa-vanya-* «vainquant tout le monde»; *aiwi-vanyah-* «qui a le dessus»; *vanō-vīspa-* «vainquant le monde»; *vanāra-* nom de démon «qui vainc les hommes»(*$*h_1$wene-$h_2$ner-o-*); *van-at-i-* «victoire», *vanat-pəšana-* «gagnant la bataille», dérivé en **-et-* comme le nom des Vénètes.

3.2. Interprétations du nom des Vénètes

Celui-ci a été diversement interprété : parfois comme «aimé», ce qui est d'autant moins vraisemblable pour un nom de peuple que l'amour indiqué par *$*h_1$wen-* est sexuel; parfois aussi par référence à v.irl. *fine*, v.isl. *vinr*, ce qui est plus grave, car comme certaines explications des noms en **s-bhe*, cette étymologie renvoie à une parenté ethnique, conçue parfois de plus comme sodalité guerrière, ce qui procède de la même conception que celle qui a fait des *ārya-* des «nobles» (cf. ἄριστος). La méthode mise en œuvre ici, et qui revient à poser que les hommes de langue indo-européenne se sont désignés eux-mêmes par des appellatifs dont ils ont fait des noms propres, par un processus bien connu dans leurs diverses langues, ne permet qu'un rapprochement : celui des formes en *$*h_1$w-en-et-* de l'avestique qui, en cette langue, sont propres à la «victoire»; celle-ci n'est que l'un des aboutissements du «désir de conquête», si bien qu'il n'est pas nécessaire

de voir dans les Vénètes des « Victorieux » : on attribuera à leur nom le sens étymologique d'où procède la valeur des termes avestiques, « désirant conquérir » ; ce sont en somme des « conquistadores », comme l'a pensé Devoto.

3.3. Extension géographique

Les lieux si nombreux auxquels les Vénètes ont laissé leur nom laissent entrevoir certaines des modalités de leur conquête.

Le nom des Vénètes peut être attesté par hitt. TUL*Wa-na-at-ti-ja-ta*, nom d'une source au Nord de la Syrie. Sinon, sa plus ancienne attestation concerne les Vénètes de Paphlagonie (dans la région d'Amisos selon Hécatée de Milet [frg. 200] *ap.*Strab. 12,830 ; et cf. Eust., *ad Il.* 2, 852), alliés des Troyens nommés peu après les Pélasges,

Iliade, 2, 851, Παφλαγόνων δ᾽ἡγεῖτο Πυλαιμένεος Ϝλάσιον κῆρ
852 ἐξ Ϝενετῶν, ὅθεν ἡμιόνων γένος ἀγροτεράων,

avec une allitération qu'offre par ailleurs le vén. *vo.l.te.r.ko.n. vo.n. ta.r.*, et qui conjoint les deux racines « désirer », *h_1w-en-* dans le nom des Vénètes, et *h_1w-l̥-* dans le syntagme dont dépend le nom du chef, à peu près « volonté » (« cœur volontaire », cf. vén. *vo.l.tiio-*) ; elle a été méconnue des sources classiques, qui ont transmis le nom sous la forme Ἐνετοί. Sous cette forme ou une autre, le nom des Vénètes se trouve dans divers toponymes et ethniques dans des conditions telles qu'ils laissent parfois entrevoir la successivité de couches indo-européennes diverses : la Baltique s'est appelée « golfe Vénète », Ptolémée 3, 5, 1, avant de s'appeler « mer suébique » ; en Sarmatie d'Europe, au bord de la Vistule, les Germains ont baptisé leurs voisins slaves du nom de *Wendes* (emprunté par les Finnois, qui ont nommé la Russie *Venät*) ; les Carpates sont appelées « mont Vénète », Ptolémée 3, 5, 4, et nommées immédiatement après le mont « Alain » ; le lac que nous nommons de Constance est *Venetus lacus* chez Pomponius Mela, 3, 2, tandis que les Allemands l'appellent « lac de Bregenz », du nom de la ville celtique auquel correspond le *Brigantinus lacus* de Pline, 9, 63 ; du Tirol à l'Alsace des montagnes et agglomérations s'appellent *Gross Venediger*, *Venetberg*, *Wennerberg, Wennederberg, Wendenheim.* Les Vénètes ont laissé une langue italique aux bords de l'Adriatique, par des inscriptions s'étageant d'environ – 550 à – 100, époque où ils se sont romanisés, et en Pannonie sur la rive droite du Danube, par des inscriptions d'une nécropole en activité entre – 440 et – 340, c'est-à-dire juste avant l'arrivée des Celtes dans la plaine danubienne ; des *Venetulani* se sont établis dans le Latium, Pline 3, 69 (tandis que dans le Sud de l'Italie, *Venusia* porte un nom apparenté) ; du côté de

l'Atlantique, ils ont laissé leur nom à *Vannes*, et un texte (à Plumergat) montre que ces Vénètes du Morbihan parlaient du gaulois ; proches, les *Veneticae insulae* (Pline 4, 109), qui ont donné lieu à des identifications variées (Belle-Ile, Noirmoutiers, Oléron...), ont dû se trouver dans la région de l'embouchure de la Loire.

3.4. Ethnographie

Cette extension géographique permet des remarques ethnographiques : les établissements vénètes se trouvent près de voies d'eau, mers (mer Noire, Baltique, Adriatique, Atlantique) et fleuves, souvent près de l'embouchure de ces derniers (Vistule, Rhône formant une sorte de delta dans le lac de Constance, Danube près du confluent de la Drave, Adige, Atlantique), et/ou de régions montagneuses pouvant fournir du bois de construction. Cela donne une coloration particulière aux « migrations » par lesquelles ces hommes ont « désiré conquérir » : l'impulsion de celles-ci est le commerce ; et les Vénètes apparaissent déjà comme des spécialistes du transport et de la circulation dans le si suggestif distique homérique cité : ils y sont caractérisés de ce point de vue aux plans de la culture et de la nature, par leurs mules, animal de transport par excellence dans l'Antiquité, et par la Paphlagonie, région côtière et montagneuse connue également dans l'Antiquité pour l'excellence de ses bois, employés pour la charpente des navires, et, pourquoi pas, pour la fabrication des chariots tirés par les mules.

Concernant les instruments de transport, animaux attelés à des chariots et chars (de transport, non de guerre) et navires, nous savons que, spécialistes des mules à l'époque de la guerre de Troie en Paphlagonie, ils le furent dans la région de l'Adriatique à l'époque de Denys I[er], tyran de Syracuse puisqu'à sa demande ils exportèrent en Sicile des chevaux (Strabon 5, 212) ; l'élevage que les Vénètes faisaient de chevaux était si célèbre qu'il est le sujet d'un mythe (Strabon 5, 215) ; ces chevaux devaient être comparables à ceux des Sigynnes, voisins des « Enètes de l'Adriatique » selon Hérodote, 5, 9 : « attelés à des chars, ils sont très vifs ; c'est pourquoi les gens du pays circulent sur des chars » (traduction Legrand). Quant aux navires, nous ne savons pas si ces Vénètes de l'Adriatique, établis dans une région où coulent l'Adige, la Gail (affluent de droite de la Drave), où le Timave prend sa source, et qui est bordée par une côte, en ont construit, et, si oui, s'ils les ont transportés sur leurs épaules pour franchir le Brenner ainsi que l'ont fait leurs voisins Istri (Plin. 3, 127-8 ; Flor. 32, 3, 13) ; mais nous savons bien quel mal les bateaux des Vénètes du Morbihan ont donné à César, qui les décrit minutieusement, *B.G.*, 3, 8 ; 13.

Concernant les voies de circulation, les Vénètes se sont établis près de routes commerciales qui leur préexistaient, par un processus qui n'a rien que de normal, et qu'illustre par exemple à l'autre extrémité du monde indo-européen, bien plus tard au VII[e] siècle, le fait que les documents tokhariens ont été trouvés en des lieux qui jalonnent une partie de la route de la soie. Ces routes, préhistoriques dans le cas des Vénètes, sont celles de l'ambre et de l'étain (deux produits dont la circulation est en partie liée, puisque l'ambre a initialement servi de monnaie d'échange contre des métaux). C'est l'ambre qui a pu attirer les Vénètes du côté de la Baltique, immense fournisseur d'ambre, et le plus ancien; et ce sont les routes de l'ambre qui suivent les voies naturelles que sont les grands fleuves de l'Europe centrale qu'ont dû emprunter ceux des Vénètes qui ont laissé leur nom à la Vistule, aux Carpates, au lac de Constance, et qui se sont établis dans l'Adriatique : dans cette région se situe un mythe d'ambre – le mythe des Héliades transformées en peupliers et dont les larmes devinrent des gouttes d'ambre lorsqu'elles pleurèrent la mort de leur frère Phaéthon, tombé dans l'Eridan, fleuve qui peut représenter en partie le Pô, mais dont le nom rappelle ceux du Rhône, Rhodanos, et de la Radaunè qui se jette dans la Baltique –; et ce mythe a été confirmé par de nombreuses trouvailles archéologiques comme celles de Frattesina di Fratta Polesine. Il n'est pas impossible que, bien plus tard, les Vénètes de l'Adriatique aient été attirés en Pannonie par le commerce de l'ambre; en tout cas Pline, 37, 44, dit que l'ambre «est apporté par les Germains principalement dans la province de Pannonie, d'où les Vénètes, que les Grecs ont appelés Enètes, ont été les premiers à le faire connaître, étant proches voisins de la Pannonie et vivant autour de la mer Adriatique» (traduction de Saint-Denis). Et c'est l'étain qui a pu attirer les Vénètes de l'Atlantique, et son commerce leur conférer une puissance si redoutable pour César.

4. *VER SACRUM* ET INTRODUCTION AUX ETHNIQUES FONDÉS SUR DES LEXICALISATIONS PRONOMINALES

Tous les hommes de langue indo-européenne ne se sont pas répandus selon les voies des Vénètes, non nécessairement belliqueuses. Moins pacifique, parfois, a dû être la fondation de colonies consécutive à une disette, ainsi dans le cas de Bellovèse, dont le départ fut à l'origine de la descente des Gaulois sur Rome selon Tite-Live, 5, 34; ou de celui du *uēr sacrum* italique par lequel, en cas d'épidémie ou de disette, étaient voués à Mars les enfants à naître au printemps suivant; parvenus à l'âge

d'homme, ceux-ci devaient partir fonder une colonie ; et l'on peut imaginer que leurs expéditions ont pu s'accompagner d'actes tels que celui que le mythe romain narre sous le nom d'enlèvement des Sabines. Quoi qu'il en soit, le *uēr sacrum* désigne pratiquement en Italie les migrations des Sabins et Samnites qui, par leur nom, nous mènent vers un ensemble d'ethniques autre que celui qui a été étudié jusqu'ici.

Ces autres ethniques sont tirés de formations pronominales figurant dans le lexique comme pronoms proprement dits, ou comme lexicalisations pronominales, ainsi que de dérivés d'appartenance-détermination de ces pronoms. Le sens de l'ethnique se tire clairement de l'appellatif correspondant, comme dans le cas des ethniques tirés de racines verbales. A la différence de ce qui se passe pour ces derniers, leur champ sémantique est structuré par une opposition qui découle du fait que, de manière générale, les pronoms sont des indices de relation qui, à ce titre, impliquent des oppositions binaires (par exemple « ici » et « là » dans l'espace). L'opposition est ici celle de l'« autre » et du « soi » : les hommes de langue indo-européenne se sont donné des noms se référant à leur qualité d'hommes venant d'un « autre » territoire, et ayant fait « leur » le territoire sur lequel ils se sont établis, ce qui correspond naturellement à une réalité historique que nous connaissons bien.

Pour ce qui est de l'« autre », c'est-à-dire des noms du type *Aryens*, cette polarité suffirait à en établir le sens ; on répondra cependant à l'objection parfois faite qu'on ne s'appelle pas soi-même « autre », que c'est pourtant une vérité d'expérience dans les minorités, qui se justifie, au plan de la langue, par la valeur inclusive de « autre » (cf. « nous autres », expression employée, par exemple, par les Juifs qui ne voulaient pas que leur qualité de « Juif » puisse être explicitement nommée, donc facilement reconnue, au moment du nazisme).

5. L'AUTRE, *ĀRYA-*

Le nom des *ā́rya-*, qui a désigné à l'époque de leur communauté les Indo-Iraniens (et dont l'Iran a tiré son nom), est apparenté en sanskrit même à des appellatifs qui sont dans le même rapport formel que par exemple lat. *ali-* et *ali-o-* : *arí-*, qui s'applique à toutes sortes de relations d'altérité, bonnes ou mauvaises, « ami » et « ennemi », « hôte », « rival », moitié exogamique dans le vocabulaire du mariage, etc. ; *árya-* et *aryà-* « hôte » et « maître de maison » (d'où, dénominatif, *Arya-mán-*, dieu du mariage ; de ceux qui se reconnaissent *ā́rya-* par opposition aux

Barbares, etc.); le nom des *Alains* (᾿Αλανοί, etc.) est, lui, une lexicalisation onomastique de skr. *áraṇa-* «étranger» (itération pronominale brisée agglutinant deux noms de l'«autre», **al-* et **-ano-*); peuvent lui être apparentés l'ethnique *Alināsaḥ*, s'il repose sur **ali-no-* avec une formation comparable à celle de celt. *Suer-āres* (*-eh_2-$s^{o}/_{i}$-), et le nom de ville *Alakā*, morphologiquement apparenté à tokh. A *ālak*/ B *alyek* «autre», et, avec d'autres valeurs relevant du champ de l'altérité, got. *alakjo* «ensemble», v. irl. *aire*, gén. *airech* «libre». Il est possible, mais indémontrable, que certains toponymes anatoliens aient le même origine : ^{URU}*Ala*, ^{URU}*Alalija* (**al-ali-*, cf. lat. *al-ter* et *ali-quis*?), ^{URU}*Alana*, à côté du nom des Alains, *Ala-sia*, nom de Chypre, qui pourrait être un dérivé d'appartenance-détermination du type *Venu-sia* sur *Ala-* < **alo-*, cf. *ali-* dans gaul. *Ali-sia*; etc. Il est sûr en tout cas que l'altérité doit être comprise comme celle du territoire, comme le montrent, sinon ces toponymes, du moins l'ethnique gaul. *Allo-broges*, «qui viennent d'une autre marche», d'un type morphologique différent et plus récent que le type *ā́rya-* : en celui-ci sont conjoints deux sèmes, l'un d'altérité, l'autre d'espace; ces deux thèmes sont dissociés et chacun d'eux précisé par chacun des deux membres du composé nominal *Allo-broges*, par un mécanisme qu'illustrent d'autres exemples, ainsi le nom du *Sam-land* à côté de l'ethnique prussien *Sémbai*; ou *Opi-tergium* (mod. *Oderzo*) «sur/vers le marché», à côté de l'ethnique *Opi-cī* «Osques» (cf. *Op-scī*). Le nom des *Allo-broges* est fondé, comme celui des *ā́rya*, sur un terme du lexique, comme lui composé de «autre» et du nom de la «marche» **mr-og-*, cf. gall. *allfro* «étranger», et avec pour nom de l'«autre» non plus *allo-*<**alno-* (à côté de **al-ano->áraṇa-*, et peut-être **ali-no-* dans *Alināsaḥ*), mais **ali-(o)-* dans run. norv. (vers + 400) *alja-markiR*. Par leur nom, les *Alsaciens* (lat. du Moyen Age *Ali-sātia*, v.h.a. *Eli-sâẓẓo*) viennent d'un «autre établissement», de l'autre côté du Rhin; et cf. v.h.a. *ali-*, *eli-lenti*, v.irl. *aili-thir*, gall. *all-myr* «d'une autre terre», «d'une autre mer»; etc.

6. Le soi

D'autres noms sont bâtis sur le réfléchi. L'opposition entre l'«autre» et le «soi» peut prendre des valeurs diverses, en partie en fonction du déroulement de la conquête : l'«autre» peut être l'«étranger» que le conquérant trouve à son arrivée, en valeur non inclusive, et en valeur inclusive, le même que le soi, s'agissant du territoire «autre» au départ, «sien» à l'arrivée. Mais l'opposition de l'«autre» et du «soi» est au

terme de la conquête neutralisée, car elle est transcendée par l'organisation politique d'une fédération (qui peut avoir pour signifiant **sem-no-*), d'une « totalité politique » (qui a pour signifiant spécialisé **teutā*). De cette nécessaire complémentarité entre le « soi » et l'« autre » témoigne l'emploi de la particule *san* accolée au nom des Barbares hitt. LU^{MEŠ}*Lulahhias-san* (cf. louv. ^{d}*Lulahinzas-tar*, et pour l'agglutination des deux particules, skr. *sanu-tár* « au loin »), d'un **se-n* qui a donné, entre autres, le nom des *Sénons*.

6.1. *Nitiobroges* et *Allobroges*

Parce que les pronoms sont asémantiques et qu'une même fonction peut être assumée par des pronoms différents (*e.g.* lat. *quī* et gr. ὅς, relatifs), les noms de l'« autre » peuvent présenter des variations telles que **ali-(o)-*, **alno-*, etc., et les noms du « soi » différer autant l'un de l'autre que par exemple **nítyo-* (skr. *nítya-* « son propre », got. *niþjis* lexicalisé pour le « parent (par alliance) », etc.) et **sĕ*, **s-wĕ*, etc. Les deux réfléchis s'opposent à des formes de **al-* « autre » dans des textes, *e.g.* pour le premier *R.V.* 3, 53, 24 ; 5, 85, 7 ; 7, 4, 7 (*áraṇa-* ≠ *nítya-*), ou pour le second *R.V.* 6, 75, 19 ; 10, 124, 2 (*áraṇa-*≠ *sva-*), v. angl. *Aelþeódige maen... swǣse maen in leódum.* A skr. *áraṇa-* ≠ *nítya-* répond dans l'onomastique gauloise *Allo-broges / Nitio-broges* (tandis que les *Combroges* [gall. *Cymmry*] sont ceux « qui ont une marche commune, sont limitrophes »). A skr. *áraṇa-* ≠ *sva-* correspondent, à côté des noms du type *ā́rya-*, des noms en **s(w)e-*.

6.2. Un nom en **s(w)e* hors de l'Europe : la *Chorasmie*

Un problème concerne l'extension dialectale de ces derniers. En tout état de cause, on ne retiendra pas les toponymes de l'anatolien (où par ailleurs le réfléchi est du thème **te*, louv. *-ti*, hitt. *-za<*-ti-o*) comme *Salma* (cf. tokh. B *solme* ? ?), *Sapana* (**sebh-°no-* ? ?). Mais alors que ces ethniques sont essentiellement européens, on peut se demander si la situation offerte par *ā́rya- / Allobroges*, *Alsaciens* (noms plus vivants en indo-iranien que dans les langues européennes) ne se reflète pas, inversée, pour les noms en **s(w)e*, puisque la Perse a au moins un exemple de ceux-ci à ma connaissance : non pas le nom de la Susiane, *huvaǰa-*, qui est à lire *Ūvǰa,-*, mais celui de la Chorasmie, *huvārazmĭ-*. C'est en effet un composé du réfléchi et d'un second membre qui, en tant que dérivé en

**-m-*, répond en iranien à l'av. *rasman-*, nom d'une unité militaire « regimen », mais qui, sans ce suffixe, répond à skr. *sva-rā́j-* « qui domine par lui-même », d'où *sva-rā́jya-* « domination propre ». Le toponyme répond alors à l'un des vieux noms de la « Suède », *Swe-rīke* (§ 6.3), l'isoglosse irano-germanique étant soulignée par l'existence, dans un autre vocabulaire, d'un composé du réfléchi s'appliquant à celui qui est mort de sa « propre » mort, mort naturelle, v.p. *uvā-maršiyu-*, v. isl. *sví-dauđr*, à second membre renouvelé et premier membre identique aux diverses formes du nom des « Suédois ».

6.3. **swī* et le nom des *Suédois*

La forme de ces noms montre bien que le réfléchi est employé pour l'appropriation du territoire : **swī*, identique au génitif du possessif du réfléchi lat. *suī*, cf. skr. *svī-kṛ-* « s'approprier ». Cette forme (qui apparaît aussi dans le nom de la *Suisse*, ci-dessous) a donné *Sviár*, mod. *Svear* « Suédois », formellement identique à skr. *svīya-* (**swī-o-*), et recaractérisé par un suffixe d'ethnique en *Suīōnēs* (Tacite), *Sueōnēs* (Adam de Brême), v. angl. *Swéhon*, etc. Par le processus qui a abouti à la formation de *Allo-brogēs* à côté de *ā́rya-*, la lexicalisation pronominale a été refaite en composés nominaux : v. suéd. *Swē-rīke* > mod. *Sverige*; *Sví-þjóđ* (run.s. *sviþiuþiu*, *sveþiuþ*). S'il ne s'est pas développé indépendamment du nom de la *Chorasmie*, le premier est un héritage, et le second en est une adaptation au système de la **teutā* (c'est l'inverse dans le cas contraire). De toute manière, le second repose simultanément sur le vieux syntagme représenté aussi par le sud-picénien SAFINA TOÚTA « totalité politique de ceux qui ont fait leur le territoire », en lequel sont dissociés les sèmes de « réunion » et « possession », comme d'une autre façon dans les noms des *Samnites / Semnons* et *Sabins/Suèves* (ci-dessous).

Ce composé souligne le caractère de l'appropriation : politique (et non politique et territoriale comme dans le composé à second membre nom du « royaume », ni seulement territoriale comme dans le cas des *Allobroges*). Il serait en tout cas d'autant plus erroné de rechercher dans ce nom, et dans d'autres, apparentés (comme **s-bhe*, ci-dessous), une désignation de parenté ethnique, raciale, comme cela a été de mise trop longtemps, que là où **swī* figure dans un nom de parenté, celle-ci est une parenté par alliance : v.h.a. *gi-swīo* « beau-frère », mari de la « sœur », cf. les noms de « beaux-frères dont les femmes sont sœurs », lit. *sváinis*, lett. *svaĩnis*, de **swoi-ni-*, sur la forme de datif du réfléchi à côté de **sei-no-* > got. *seins*, etc., adjectif possessif, et de gr. εἰλίονες < **swei-li-* + *-on-*, αἴλιοι, ἀέλιοι < **sm̥-swi-*, **sm̥-swe-li-o-*, etc.. *Svíþióđ* s'oppose à gall. *all-tud*, v. angl.

aelꝥeód « d'une autre **teutā́* » comme *Nitio-broges*, skr. *nítya-* et *sva-*, à *áraṇa-*.

6.4. **s(w)e-t...* : Suetidī, Svēte, Suisse, Sētia, Saticula, Satricum, etc.

C'est à des agglutinations pronominales qu'appartiennent toutes les autres formes sous réserve peut-être de certaines formes en **s(w)e-t...*, dont la dentale peut être un suffixe nominal et/ou un thème pronominal comme le second constituant de *ní-tya-* (apparenté à la particule réfléchie hitt. *-za*). Ces formes sont les moins claires de toutes celles qui seront examinées, pour cette raison, et parce qu'on a fait de certaines d'entre elles des hypocoristiques du composé *Sví-ꝥióđ* (alors à mettre au compte des désignations relevant de **teutā*). C'est le cas pour *Suetidi*, d'où peut-être le nom du dieu Óđinn, *Sviđrir*, *Sviđrur*, *Sviđurr*, si on le comprend comme « dieu des Suédois ». Mais, parmi les autres, il en est qui peuvent reposer sur **s(w)ĕ̄-ti-* (agglutination de deux thèmes dont chacun a pu donner un réfléchi), cf. alb. *ve-të* « soi-même » < **swe-ti-*, et lit. *svetỹs*, *svẽčias* « hôte », v.sl. **sětŭ* dans sl. *po-sětiti* « inviter » : ainsi le toponyme lette *Svēte*; le nom de la *Suisse* /*swīts*/ *Schwyz*, à l'origine nom d'un village (entre les lacs de Zug et des Quatre-Cantons), puis de son canton, ensuite employé par les Autrichiens pour la confédération, sous forme *Schweiz* (avec **ī* > *ei* dans les parlers bavarois de l'Autriche), qui peut reposer sur **swīt(a)-*; et les villes d'Italie *Sētia*, cf. **sēto-* > v.sl. **sětŭ*, et avec le même vocalisme que dans le nom des Sabins et Samnites, probablement osco-ombrien, **s°-* (forme avocalique du pronom, et anaptyxe), *Saticula*; et *Sa-tri-cum*. Ce dernier doit avoir le suffixe nominal d'opposition (cf. lat. *cēterī* « les autres » < **ke-e-tero-*, et les toponymes v. irl. *Airthir* « qui est devant », *air-* < **pari-*, ou *Sues-tra*, ci-dessous); pour le sens, *Satricum* peut désigner le « soi » comme *Sabīnī*, ou la « réunion » comme *Samnītēs* : cf. skr. *satrā́*, av. *haθrā̆* « ensemble » (de la même forme **s°-tr-* ou de **sm̥-tr-*)).

Sous bénéfice d'inventaire, on ajoutera aux formes en dentale le mod. *Saix* (en Poitou) < **Suēdas*, *Suaedas*, dont le *-d-* peut venir de la sonorisation romane de **-t-* ou être un ancien **-d-*, cf. ἴδιος « propre à » < **swi-di-o-*; ambiguïté du même genre dans l'ethnique germanique transmis par Ptolémée, 2, 11, 14, Σειδινοί (*v.l.* Σιδεινοί), dont la dentale peut être un **d* ou un **t* si la forme a été soumise à l'action de la loi de Verner.

6.5. _*s(w)e-(no-, -ro-)_ : _Sénons, Suerārēs_

Comme les noms de la Suisse et des Suédois, d'autres ethniques sont en rapport avec des possessifs et réfléchis (de cas direct ou indirect comme _*s(w)e_ et _*s(w)_$^{e}/_{o}$_i_). Ainsi les noms des _Sĕnōnēs_ « l'un des peuples gaulois les plus puissants », Cés., _B.G._ 54, 2 (auxquels on attribue la prise de Rome vers – 385) ; de l'île de _Sein_ ; des _Suen-cī_ (même suffixe que dans _Opicī_), d'où _Sainex_ (dans la Nièvre) : cf. outre la particule d'altérité _-san_ ajoutée en hittite au nom des « Barbares », de _*s(w)e-n_ $^{i}/_{o}$- le réfléchi tokh.A _ṣñi_/B _ṣañ_ et son dérivé _ṣaññe_ « en propre », ainsi que AB _ṣñaṣṣe_ « parent (par alliance) », et de _*sw_$^{e}/_{o}$_i-n_$^{i}/_{o}$-lit. _svaínis_, etc. (et v. sl. _*svěnŭ_ dans _svěnje_ « à l'écart de », _o-svěniti_ « se tenir à l'écart », avec sèmes d'altérité et d'espace) ; parallèle peut être, en _*swe-r(o)-_, _Suerārēs_ (_Sougères_ sur Yonne), cf. _*swo-ro-_ dans alb. _vajë_ < _*swariā_ « jeune fille », _*sew(_$^{e}/_{o}$_)-ro-_ dans arm. _iwr_ « sui, sibi ».

6.6. _*s(w)es-so-_ : _Suessa, Suessula, Suessionēs_

Certaines formes sont faites sur le réfléchi et/ou son adjectif au datif (cas qui prédique la possession, à côté du génitif qui prédique l'appartenance), comme _*s(w)es-so_, cf. ombr. _sueso_ « suō », _seso_ et got. _sis_ « sibi », et, lexicalisés, got. _swēs_ « en propre », v.h.a. v.s. _swās_ « en propre, aimé », v.fr. _swēs_ « apparenté » ; en Italie, _Suessa_, nom de deux villes distinguées par leurs épithètes, _Suessa Aurunca_ (mod. _Sezza_), ville principale des _Auruncī_, et _Suessa Pōmetia_ (qualificatif latin par opposition à _Abella_), et les deux _Suessula_, l'une en Sabine au voisinage de Reate, l'autre en Campanie ; chez les Germains, _Suesia palus_, nommé comme l'un de leurs plus grands marécages, Mela 3, 29 ; chez les Celtes de la Gallia Belgica, les habitants du _Soissonnais_, _Suessiōnēs_, et chez les Celtes d'Espagne les _Suess-et-ānī_ (deux suffixes d'ethniques) ; avec des suffixes en _*-t-_, _Sues-tra_, mod. _Susteren_ dans la province du Limbourg, _Suess-ātiō_, Itin. 454, 9 ; Σουεστάσιον (_Sues-tāt-io-_), Ptol. 2, 6, 64.

6.7. _Sabins, Suèves et Samnites, Semnons_

Apparentées également à des réfléchis et possessifs sont les formes en _*s-bhe_, cf. gr. σφέ, σφός, et _*s(w)ē-bh-_, cf. lit. _sēbinu_, _sēbinti_ « mettre

de côté» : de **s°-bho-*, **Sabus* dans les anthroponymes *Sabō*, *Sabeius* (**s°bhei-o-*, cf. lat. *sibī*, et pour la structure lat. *meus*, skr. *svay-á-*) d'Italie, et le théonyme *Sābus* bâti sur lui par allongement de la voyelle radicale («*vr̥ddhi*») comme *ā́rya-* sur *árya-*, et dans le dérivé d'appartenance (*-ī-*) – détermination (*-no-*) *Sabīnī*, avec **bh* > lat. *b*, o.o. *f* (SAFINA TOÚTA, etc.) ; de **swē-bho-*, le nom des *Suēbī* transmis par Tacite, cf. v.h.a. *Swāba*, v.isl. *Sváfa-(land)*, run. suéd. *swāƀa-harjaR* ; etc.

6.7.1. Problèmes sémantiques

Comme le nom des Suédois, cet ethnique permet de bien poser le problème sémantique : «possession» ou «parenté (ethnique)» ? C'est le plus souvent par le second de ces sens qu'on explique ces noms, avec une référence particulière à la famille d'all. *Sippe*, au mépris de plusieurs faits : la parenté étymologique, pourtant simultanément admise, avec skr. *sabhā́* «réunion (sociale»), qui implique que *Sippe* s'applique non à la parenté par le sang, mais à la «réunion (familiale)» des deux types de parents, par le sang et par alliance (évoquée ici à maintes reprises par des noms de «beaux-frères»), «réunion» qui caractérise tous ceux que la «grande famille» considère comme «siens», **s(w)e* (comme dans *sabhā́*, *Sippe*), au terme de l'échange exogamique ; le fait que ces noms sont, comme ceux qui les ont précédés dans cet exposé, en rapport avec des possessifs et des réfléchis, ce que soulignent les lexicalisations v.sl. *sobiti* «s'approprier», *sobina* (même forme que *Sabīno-*) «possession». Le problème sémantique se déplace alors du dilemme «possession»/«parenté ethnique» au dilemme «possession»/«réunion». La diversité des couches d'hommes qui ont apporté en Italie les divers parlers osco-ombriens (qui peuvent remonter au «sabin»), et la diversité des populations suèves énumérées par Tacite (*Germ.* 38-44) donnent à penser que les noms des Sabins/Suèves a pu s'appliquer à la «réunion» de ceux qui ont fait «leur» le territoire ; mais en fait les deux sèmes, dont rien n'empêche la conjonction en une lexicalisation pronominale, ont été répartis en deux formes, toutes proportions gardées comme en *Suīpiōđ/safīna toúta* : le nom des Sabins/Suèves a été restreint à la valeur de «possession» par suite de l'emploi pour des rameaux qui se sont détachés d'eux de noms à valeur de «réunion», comme le thème **sem-* dont ils sont des adjectifs déterminés, ceux des *Samnites* (ethnique dérivé du toponyme *Samnium*, à même suffixe que par exemple tokh. B *ṣaññe*) et *Semnons* ; en **s($^{e}/_{o}$)m-n $^{i}/_{o}$-*, ils sont étymologiquement apparentés à des formes qui désignent dans le lexique la «réunion», de l'Inde (skr. *sámana-* «ensemble») à l'Irlande (v. irl. *samain*, nom d'une fête fédérale annuelle), en passant par le grec (ἅμα < **s°m-n̥*, avec anaptyxe comme dans *Samn-*), le germanique (v.h.a. *samana* «ensemble», all. *zu-sammen*), etc.

Par leur nom, les Samnites/Semnons sont désignés comme « réunis » aux Sabins/Suèves.

6.7.2. Données culturelles

On ne peut en effet considérer ces deux paires de noms comme indépendantes l'une de l'autre : la hiérarchie Sabins/Suèves – Samnites/Semnons est un héritage culturel et onomastique commun à ceux des Italiques que nous appelons Osco-Ombriens, et dont le plus ancien parler est le « sabin », et aux Germains que décrit Tacite, en les enfermant par un procédé de composition circulaire entre ceux de leurs noms qui sont faits sur le thème du réfléchi, *Suēbī* (suivi de *Semnōnēs*) et *Suīōnēs* : insistant d'abord sur la diversité des *Suēbī* (38), il passe aux *Semnōnēs* (39), « *vetustissimos... nobilissimosque Sueborum* », puis aux nombreux autres peuples suèves (40-43), aux noms sans rapport avec le réfléchi, pour finir par *Suīonēs* (il parachève ensuite sa description par la mention des peuples non germaniques des bords de la mer *suēbicum* [auparavant « vénète » = Baltique], 45, et conclut, 46, « *Hic Suebiae finis* », par une composition circulaire complète). Or nous savons que les *Samnites* sont un rameau des *Sabins* par exemple par Strabon (5, 4, 12) qui, dans le même ordre de composition que Tacite, énumère ensuite des peuplades de moindre importance également aux noms divers, essentiellement *Hirpīnī* et *Pīcentēs* (5, 4, 13). Dans une étude plus poussée, ceux-ci, reliés à la louve et au pic de Mars, dieu du *uēr sacrum*, apparaîtraient comme faisant partie d'un ensemble structuré en fonction de ce dernier : *Marsī, Marrūcīnī, Mamertīnī* : Mars; *Frentānī*, sur un nom du « cerf », animal dont la ramure symbolise le printemps animal et végétal, et *Sacrānī* : uēr sacrum; *Pīcentēs* et *Hirpīnī*, symboles de la conception animale et de la nourriture des petits mammifères (ainsi que de l'agriculture : cf. *hirpex*), tandis que *Bouiānum* l'est de l'économie tout entière, pastorale et agricole; etc.

6.8. Autres formes en rapport avec **s-bhe : Sémbai, Sabellī, Suebrī, Serbes***, etc.

Les thèmes des noms des Samnites/Semnons et des Sabins/Suèves sont entrés en collision dans les formes en **sem-bh-* que sont d'une part l'ethnique de Prusse *Sémbai* (*Sembī*, Adam de Brême, *Sambītae*, Peter von Duisburg) et le toponyme *Sam-land* < **semb-land*, et d'autre part lett. *sę̄brs* « ami » à côté de *sebris*, ci-dessous. Aux noms des Sabins/Suèves sont apparentées des formes complexes qui peuvent avoir un rapport plus ou moins lointain avec un génitif pronominal, un possessif et/ou un réfléchi :

1) **s°bhe-l-no-* dans le nom des *Sabellī* osco-ombriens, avec les mêmes constituants mais dans un ordre différent, que dans **s(w)e-l-bho-* >angl. *self*, etc. ; c'est un adjectif déterminé en **-no-* bâti sur ce qui serait un génitif pronominal en **-l-* dans une langue anatolienne, cf. hitt. *sel*, anaphorique, *apel*, démonstratif éloigné, de **bhe-l* à côté de **bhi-l(-i-, -o-)* dans lyd. *bilis* « son propre », gr. φίλ(o)-, indiquant la possession inaliénable, et lexicalisé pour l'« ami » ;

2) **s(w)e-bh-ro-* (agglutination qui en ses derniers éléments rappelle (ὄ)-φρα, etc.) dans les ethniques celtes *In-subrī* (cf. pour **su-bho-* v.pr. *sups* « son propre »), *Suebrī* (cf. pour **sĕ̄-bh-ro-* lit. *sēbras*, lett. *sebris* « ami », etc.) ; un adjectif déterminé est tiré de cette forme dans l'anthroponymie d'Italie, **s°bh-r* $^{i}/_{o}$*-ni-o-* > *Sabernius*, *Safrōnius* ;

3) **s(w)e-r-bho-* qui est à **s(w)e-bh-ro-* ce que **s(w)e-l-bho-* est à **s(we)-bhe-l-(no)-*, dans le nom des slaves *Serbes, Sorbes*, auquel correspond dans le lexique russe *pa-serb* « beau-fils » ; alors que **s°bhe-l-(no)-* a un parallèle dans hitt. *apel*, **s(w)e-r-(bho)-* en a un dans le nom des *Suerārēs*, et les possessifs correspondants.

On ne peut dire si ces noms ont la seule valeur de « possession » des ethniques en **s(w)ē-bho-*, ou, de plus, la valeur de « réunion » qui y est tombée en déshérence par suite de l'emploi, à côté de *Sabīnī*, *Suēuī*, de *Samnītēs*, *Semnōnēs*. Mais c'est la valeur de « réunion » qu'offrent tous les autres des noms ici présentés, qu'ils aient quelque chose à voir étymologiquement avec **se* (auquel est apparenté **se-m*), ou non, comme en **te* (souvent associé à **se* dans d'autres champs sémantiques) les noms tirés de celui de la **teutā*, et en **alo-* « autre » et « tout », l'un des noms des Allemands.

6.9. **s(w)el-(-wo-, -no-)* : *Slaves, Hellènes*

En **s(w)e-l-*, de même structure que **s°bhe-l-* et **s(w)e-r-*, sont les noms du peuple dont les Serbes font partie, les Slaves eux-mêmes, et leurs voisins hellènes. La forme **s(w)e-l-* a pu en effet s'appliquer à la totalité : **s* $^{e}/_{o}$*l-m-* > tokh.B *solme* « tout » ; **s*$^{e}/_{o}$*l-wo-* (cf. ombr. SEVOM « totum » pour le second élément de l'agglutination) pour le « tout » avec vocalisme **-o-* (skr. *sárva-*, gr. ὅλος, gall. *holl*), et pour la « possession » avec vocalisme **-e-*, v. irl. *selb* « propriété, possession », gall. *helw*, gaul. *Lugu-selva* « propriété du dieu Lug » (et cf. *s°l-wo-* dans lat. *saluus*, pour l'intégrité physique). Sur **sol-wo-* repose le nom des *Slaves* (avec **ol* > *la* comme dans v. sl. *lani* < **olnei* « de l'an passé ») ; de même radical, le nom des *Slovènes* peut avoir le même suffixe **-ēn-* que le nom des *Hellènes*. Celui-ci entre dans une famille de mots en **s(w)e-l-*, qui peut :

avoir des vocalismes divers ; être déterminé par *-*no*-, et parfois alors passer aux thèmes sigmatiques comme Ϝhε-διέστᾱς à côté de ἴδιος ; avoir un doublet composé en **sm̥-s(w)el-* (d'où une dissimilation d'aspiration, l'esprit doux venant ailleurs d'une psilose), et être en cela comparable au nom des « beaux-frères », **sm̥-sw*$^{e}/_{i}$*li-o-* > ἀέλιοι, αἴλιοι et sans **sm̥-* v. isl. *svilar*, gr. εἰλίονες (autre composition dans le m. néerl. *swā-sel-inc* « beau-frère ») :

**sm̥-sw(e)-l-n-es-*>ἀϜλανέως, adverbe, et ἀολλής, ἀελλής « rassemblé » ;

**sw-l̥-i-* (ou **sw°-l-i-*? cf. ἅμα, et lat. *similis* < **sm̥-°li-* ou **s°m-°li-*?) > ἅλις « en masse » ;

**sw-l̥-n-es-* > ᾱ̔λής « rassemblé », et **swe-l-no-* (parallèle à **s°bhe-l-no-* > *Sabellī*) dans ἀπέλλαι « assemblées », εἰλέω dénominatif « rassembler » (d'une forme en *-*no*- ou en *-*n-es*-). C'est sur **swe-l-no-* qu'est fait Ἑλλοί et avec le suffixe d'ethniques *-*ēn*- Ἕλληνες (qui a donc un correspondant formel dans le vocabulaire de la parenté, comme le nom des *Serbes* dans russ. *pa-serb*). D'abord désignation d'une population de Dodone, le nom a été étendu à l'ensemble des Grecs, comme ceux des Francs, Alamannī, Angles ou Bretons à l'ensemble des Français, Allemands, Anglais/Britanniques.

7. Totalités transcendant l'opposition du soi et de l'autre : Teutons, Allemands

A côté de ces noms en **s(w)e* indiquant comme ceux en **sem-* une « réunion », a existé un nom de la « totalité politique », **teutā*, largement répandu dans l'onomastique (par exemple pélasgique ou illyrienne), et appellatif attesté en baltique, germanique, celtique, messapien, osco-ombrien, vénète, c'est-à-dire dans toute l'Europe sauf là où sont employées d'autres formes, ce qui est le cas dans l'onomastique pour le **s(w)* $^{e}/_{o}$*l-* (-*wo*-, -*no*-) des noms des Slaves et Hellènes, et dans le lexique en latin : la forme étymologiquement correspondante y a donné l'adjectif *tōtus*, et la forme sémantiquement correspondante y est *cīui-(tās)*, qui procède d'une autre agglutination, **kei-wi-* ; celle-ci associe non le « soi » et l'« autre » (désigné comme « lointain » dans **s-we*, comme « (l'autre de) deux » dans **s-bhe*), mais le « proche » (**kei-* de **ke*) et le « lointain » (**we*), et désigne quelque chose qui est voisin de τὸ κοινόν, Hdt. 1, 67, etc. Le nom de la **teutā* a donné en Italie les ethniques des *Praetuttii* au Nord d'Adria et des *Tūtīnī*

en Calabre, et chez les Germains, des *Teutōnēs*, *Teutōnī*, attesté pour la première fois dans le récit du voyageur Pythéas (vers – 340), et dont le dérivé adjectif (v.h.a. *diutisk*, etc.) s'est appliqué d'abord à la langue (populaire par opposition à la langue savante), avant de désigner ceux que nous appelons en français non pas *Deutschen*, mais *Allemands*.

Ce dernier n'est que l'un des renouvellements des ethniques qui indiquent qu'est transcendée l'opposition du « soi » et de l'« autre », par la « réunion » de « tous les hommes », ce que Tite-Live appelle dans le lexique latin *gēns ūniversa*, 1, 1, en décrivant les conditions d'établissement des Vénètes, en des termes qu'on pourrait probablement appliquer à toutes les couches d'hommes de langue indo-européenne venus d'« ailleurs ». Au terme de l'évolution, **al-o-*, qui donne un nom de l'« autre » dans le nom des Alains apparenté à celui des Aryens, a le sens « tout » dans *Ala-manni*, ᾿Αλαμανοί, composé à second membre nom de l'« homme » (apparenté en germanique au nom indien qui a servi de base à l'ethnique *Manuṣ* désignant les Tsiganes) ; il est du type *Allo-broges*, c'est-à-dire substitué à une plus ancienne lexicalisation pronominale onomastique ; par une réduplication, qui évoque l'itération brisée **teutā* (**te-u-*, cf. skr. *tá- u* + **-tā*; pour l'itération cf. myc. *toto* [*weto*], etc.), la *Skeireins*, 8, 12, a *allaim alemannam* « totalité de tous les hommes ».

Sans aucun doute cette organisation politique, jointe à l'immense science dont témoignent leurs poètes, a, autant et plus que le char de guerre, permis aux hommes dont la langue était indo-européenne d'imposer celle-ci sur les territoires que, venus d'ailleurs, ils firent leurs, en absorbant linguistiquement tant de leurs prédécesseurs dont, pour cette raison même, nous ignorerons toujours les langues et souvent jusqu'aux noms.

CHAPITRE 4

LE TOKHARIEN

Lambert ISEBAERT

1

Les fouilles systématiques entreprises en Asie centrale dès la fin du siècle dernier par des expéditions archéologiques allemandes, britanniques, françaises, russes et japonaises mirent au jour de nombreux documents bouddhiques, rédigés dans diverses langues. Certains textes en particulier (dont le nombre allait s'élever à plusieurs milliers au total) étaient écrits en caractères brāhmī de provenance indienne et attestaient deux formes distinctes d'une langue jusqu'alors inconnue. En 1908, celle-ci reçut le nom de « tokharien » [1] d'Emil Sieg et Wilhelm Siegling, dans un article célèbre qui peut être considéré comme l'acte de naissance de la « tokharologie » ; par une série d'étymologies évidentes, ressortant principalement au vocabulaire élémentaire (noms de parenté, numéraux, etc.), les deux savants

1. Consacrée par l'usage, la dénomination de « tokharien » (que l'on a proposé de corriger en « tougrien ») est erronée du point de vue de l'histoire : les vrais Tokhariens (ou « étéo-tokhariens ») sont des Iraniens de l'Est.

firent la démonstration irréfutable de l'origine indo-européenne des langues tokhariennes, distinguées par convention comme A et B : A *pācar*/B *pācer* « père », A *mācar*/B *mācer* « mère », A *pracar*/B *procer* « frère », A *ckācar*/B *tkācer* « fille », A *tre*/B *trai* « trois », A *ṣäk*/B *ṣkas* « six », A *ṣpät*/B *ṣukt* « sept », A *okät*/B *okt* « huit », A *känt*/B *kante* « cent », A *por*/B *puwar* « feu », A *want*/B *yente* « vent », A *mañ*/B *meñe* « mois », A *ñom*/B *ñem* « nom », etc. Progressivement, il apparut que les représentants A et B, étroitement apparentés, présupposaient un ancêtre commun, le proto-tokharien, langue indo-européenne autonome, de caractère *centum*[2].

2

Les formes littéraires les mieux représentées dans les documents tokhariens relèvent du drame, de la poésie, du récit et du genre didactique (philosophique) ; la littérature proprement scientifique comprend la médecine et la magie, la grammaire et l'orthographe. Les premiers déchiffrements furent précisément rendus possibles par la reconnaissance de nombreux parallèles littéraires dans d'autres langues bouddhiques (sanskrit, pālī, tibétain, chinois) et par l'examen des documents bilingues sanskrits-tokhariens : la comparaison méthodique avec le modèle textuel s'avérait en principe le seul guide fiable des identifications lexicales et grammaticales. Mais l'interprétation philogique des écrits tokhariens (dont l'édition se poursuit jusqu'à nos jours en Allemagne, en Chine et en France) souffrait fortement du caractère lacunaire des fragments conservés.

3

Les deux variantes linguistiques du tokharien se différencient avant tout par leur distribution géographique. Le tokharien A, appelé aussi tokharien oriental ou agnéen était le parler de la région d'Agni, à l'est du Turkestan chinois actuel, avec comme centres principaux Turfan et Karachahr ; quant à la langue B, son usage avait dû s'étendre de la région de

2. Le trait « centum » tient au fait que les anciennes occlusives labiovélaires sont restées distinctes, alors que les vélaires pures et les palatales ont fusionné.

Koutcha, située à l'ouest (d'où les dénominations de tokharien occidental ou koutchéen), vers l'aire orientale, où elle supplanta finalement l'autre langue.

De cet état de choses, on peut déduire qu'à l'époque où les textes tokhariens ont été mis par écrit, c'est-à-dire aux VIIe et VIIIe siècles de notre ère, la langue orientale n'était plus qu'un idiome religieux figé. De fait, le tokharien B, seule langue d'utilisation courante, est loin de présenter un aspect homogène et uniforme; comme tout parler vivant, il comporte plusieurs registres stylistiques et couches dialectales. Les documents tokhariens non littéraires (inscriptions et graffites, pièces de comptabilité, lettres, laissez-passer de caravanes) sont d'ailleurs exclusivement rédigés dans la langue B ; et de nombreux mots occidentaux ont pénétré par voie d'emprunt dans les textes en langue A.

A partir du IXe siècle ap. J.-C., les Turcs ouïgours s'approprièrent les territoires d'Agni et de Koutcha, mettant ainsi fin aux langues et civilisations tokhariennes, vraisemblablement éteintes avant l'an mille.

4. TYPOLOGIE GRAMMATICALE

Les fragments de langue tokharienne, « débris d'une grande ruine » (A. Meillet), ont constitué une source de progrès et de renouvellement non négligeable pour la grammaire comparée indo-européenne, malgré leur distance géographique et chronologique, ainsi par exemple dans le cas de la théorie des laryngales. Il n'empêche que les deux idiomes A et B présentent une forme déjà évoluée (dite « moyenne ») du type linguistique indo-européen, ce qui est visible surtout dans le domaine de la morphologie nominale; les modifications sensibles que le système grammatical du tokharien a subies à cet endroit doivent être attribuées au moins en partie à des influences étrangères, c'est-à-dire non indo-européennes.

4.1

Le système du genre nominal est altéré dans son principe même : comme dans certaines langues romanes, l'ancien neutre est devenu un « ambigène » (ou genre alternant), coïncidant au singulier avec le masculin, au pluriel avec le féminin. Aux nombres traditionnels singulier, duel (en voie de disparition) et pluriel, le tokharien a ajouté un « pariel » (réservé à la dualité naturelle) et un pluriel distributif.

La déclinaison du substantif distingue les cas primaires (nominatif, « cas oblique » ou accusatif et génitif), hérités comme tels de l'indo-européen, des cas secondaires (en partie divergents en A et en B), issus

de particules ou de postpositions autonomes elles-mêmes indifférentes au nombre mais agglutinées à la forme d'accusatif singulier ou pluriel. On a ainsi pour A *waṣt* « maison », plur. *waṣtu* (nom « ambigène », où nominatif et accusatif sont identiques, remontant au neutre indo-européen sg. **wāstu*, plur. **wāstu-*H_2) les exemples de flexion secondaire suivants, en l'espèce l'ablatif et le locatif :

abl. sg. *waṣt-äṣ*	abl. plur. *waṣtw-äṣ*
loc. sg. *waṣt-aṃ*	loc. plur. *waṣtw-aṃ*, etc.

Un fait intéressant laisse entrevoir encore la genèse des cas secondaires; le « cas oblique » ou accusatif qui, à l'origine, marquait non seulement l'objet direct mais aussi la direction, n'a gardé cette valeur locale que dans quelques tours où il indique le but de façon vague et générale; partout ailleurs l'*accusativus finalis* a été rendu plus précis par l'addition de postpositions (comparables à l'enclitique grec -δε), processus qui a finalement abouti à la création de l'allatif et du locatif.

À la double stratigraphie casuelle répond, au plan de la syntaxe, la « flexion de groupe » propre à l'adjectif épithète : l'accord avec le substantif fléchi à un cas secondaire se fait invariablement à l'oblique (singulier ou pluriel) : A *tsopatsäṃ* (obl.) *waṣtäṣ* (abl.) « (hors) de la grande maison » à côté de *tsopatsäṃ* (obl.) *waṣtaṃ* (loc.) « dans la grande maison ». Des terminaisons spécifiques caractérisent les noms propres par rapport aux noms communs, et à l'intérieur de ceux-ci, la classe des être doués de raison par rapport à celle des êtres non doués de raison (inanimés et désignations d'animaux).

4.1.1

L'essentiel de ce qui précède se laisse bien illustrer par les traits suivants de la déclinaison thématique en tokharien commun :

		masc./fém.			neutre
sg.	nom.	*-e < *-o-s	sg.	nom.-acc.	*-e < *-o-m
	acc.	*-e < *-o-m *-en < **-on-m̥			
	gén.	*-ense < **-o-s(y)o *-i < *-iH			
plur.	nom.	*-i < *-o-i *-eñ < **-on-es	plur.	nom.-acc.	*enta < **-ont-H
	acc.	*-ens < *o-ns			

Les formes prises par le nominatif en tokharien témoignent de la faiblesse caractéristique de la fin de mot dans une langue indo-européenne

de type médiéval. Au genre inanimé, l'altération des finales a entraîné la coïncidence des noms thématiques avec les noms sigmatiques en _*(-es-)/-os_ et avec les thèmes en _*-nt-_ (nom. -acc. _*-ont_) ; une des conséquences majeures de cette fusion morphologique est l'extension analogique de la terminaison du pluriel _*-enta_ (réservée au départ aux seuls thèmes en _*-nt-_) à l'ensemble des neutres en _*-e_ : réinterprétée en _*-e-nta_, cette finale montre bien comment un *suffixe* peut aboutir historiquement à une *désinence*. La même remarque vaut dans le cas des animés, pour les finales _*-en_ (sg. acc. des êtres doués de raison) et _*-eñ_ (plur. nom. parallèle à _*-i_ <_*o-i_, désinence d'origine pronominale) : il s'agit, ici, de l'intrusion des thèmes suffixés en _*-(e/o)n-_ à valeur individualisante dans le paradigme thématique. Le génitif singulier en _*-i_, limité aux noms de parenté et identifiable par exemple dans la forme B *āwi* « du grand-père » (indo-européen _*H$_2$ew-iH_), devait à l'origine marquer l'appartenance ; comme dans les langues italiques, il se trouve à côté du génitif thématique classique _*-o-s(y)o_, qui devait donner en tokharien, d'abord _*-ese_, puis _*-ense_ sous l'influence de l'accusatif.

4.1.2

La flexion athématique des langues tokhariennes, très précieuses à cet égard, reflète en partie les anciennes alternances vocaliques de l'indo-européen. Les propriétés caractéristiques des paradigmes indo-européens à accent mobile apparaissent clairement dans des noms comme B sg. nom. _*maṣce_, acc. *maṣc* « poing », B sg. nom. *plāce*, acc. *plāc*, plur. acc *plātäṃ* « parole, discours », B sg. nom. *yṣiye*, acc. *yaṣi* « nuit », etc., même si le vocalisme radical a été partout uniformisé par analogie : il faut partir d'une déclinaison de type « kinétique » à l'origine (sg. nom. _*mus-tḗi-s_, acc. *méus-ti-m* ; sg. nom. _*bhl̥H$_1$-tḗu-s_, acc _*bhélH$_1$-tu-m_, plur. acc. _*bhélH$_1$-tu-ns_ ; sg. nom. _*H$_2$us-i(y)ḗH-s_, acc. _*H$_2$wés-iH-m̥_, etc.). Cette alternance entre nominatif et accusatif s'observe fréquemment dans les noms féminins en laryngale, cf. B. sg. nom. *kantwo*, acc. *kantwa* « langue » < indo-européen _*dn̥g'hw-éH$_2$_, _*dn̥g'hu-H$_2$-m̥_ (forme normalisée), et inversement, B sg. nom. *śana*, acc *śano* « épouse » < indo-européen _*g^{w}énH$_2$_, _*g^{w}eneH$_2$-m_ (forme normalisée pour _*g^{w}néH$_2$-m̥_), B sg. nom. *ṣarya*, acc. *ṣaryo* « chérie, bien-aimée » < indo-européen _*swésr-iH$_2$_, _*swesr-i(y)eH$_2$-m̥_ « sororale » (forme normalisée), etc.

Du type indo-européen de flexion « acrostatique », le koutchéen et l'agnéen conservent des traces également intéressantes : la forme, récemment découverte, de B plur. nom. *awi* « brebis » (_*āwí(y)äs_ < indo-européen _*awi-es_) ne peut se comprendre que si l'on remonte à un ancien paradigme constitué de deux allomorphes indo-européens _*H$_2$ówi-_ > _*owi-_ (thème fort) et _*H$_2$éwi-_ > _*awi-_ (thème faible), fonctionnant sur le modèle exact de _*póti_ : _*péti-_ « seigneur », _*(H)ógni-_ : _*(H)égni-_ « feu », etc. De la même

façon, les substantifs A *ñom*, B *ñem* « nom », (A)B *yerpe* « disque », A *śanw-eṃ* « mâchoire » doivent provenir de thèmes forts à vocalisme radical allongé indo-européen **H_1nḗH_3-mn̥*, **Ḣ$_3$ḗrbh-i-*, **ǵḗn-u-*.

4.1.3

Si la composition nominale s'est fortement développée par imitation du sanskrit, l'union asyndétique du type A *ak-mal* « visage » (reposant sur les termes complémentaires « œil » et « nez ») ou du type A *ñom-klyu*, B *ñem-kälywe* « renom » (formé à partir des [quasi-]synonymes « nom » et « gloire ») pourrait continuer un modèle de phraséologie indo-européenne. Par ailleurs, la concurrence du génitif adnominal et de l'adjectif de relation est un phénomène également hérité.

4.2

Le système de la conjugaison tokharienne, construit sur trois thèmes principaux (le présent, le prétérit et le subjonctif), ainsi que le fonctionnement de la syntaxe verbale ont gardé l'essentiel du type indo-européen. Dans certains cas cependant, des distinctions formelles héritées comme telles de la langue-mère ne répondent plus à une fonction spécifique : ainsi la voix moyenne a largement perdu sa valeur propre pour n'apparaître plus guère qu'à titre de variante stylistique de l'actif. Comme dans les langues indiennes, beaucoup de verbes simples sont doublés d'un thème causatif : ici encore, le sens du thème collatéral s'est souvent fortement affaibli. Les grandes règles de la syntaxe (et de la stylistique) verbale demeurent familières : si l'impératif, par exemple, est le mode normal pour l'expression de l'ordre (à côté de l'infinitif « jussif »), il est exclu de la défense, où apparaissent l'indicatif et le subjonctif; dans le récit apparaît souvent, à côté du prétérit et de l'imparfait, le présent historique pour mettre en lumière les moments décisifs de l'action. La position normale du verbe conjugué est, pour les textes en prose à tout le moins, en fin de proposition. D'un radical donné s'obtiennent, par dérivation régulière, les participes et plusieurs sortes d'adjectifs et de substantifs verbaux. L'absolutif, construit sur l'ablatif du nom verbal, s'est largement diffusé sous l'influence du sanskrit.

Nous donnons ci-dessous, pour illustrer le développement de quelques finales caractéristiques (nées plus d'une fois de la désinence augmentée d'une particule déictique), le paradigme tokharien de l'ancien présent thématique *āk-* « conduire » < indo-européen **ag'-e/o-* :

agnéen	koutchéen
ākam < **ag'-o-mi*	*ākau* < **ag'-ō* (+ **w-* ?)
āśt < **ag'-e-tH_2e*	*āśt* < **ag'-e-tH_2e*
āśäṣ < **ag'-e* (+ **si* ?)	*āśäṃ* < **ag'-e* (+ **nu* ?)
ākamäs < **ag'-o-mes* (+...)	*ākem* < **ag'-o-mes*

āśäc < **ag'-e-te*	*āścer* < **ag'-e-te* (+...)
ākeñc < **ag'-o-nti*	*ākeṃ* < **ag'-o-nt*.

5. Le « proto-tokharien »

La reconstitution d'un état de langue prototokharien, outre qu'elle permet de mieux appréhender les tendances générales de l'évolution linguistique de l'agnéen et du koutchéen, est en fait la condition *sine qua non* pour confronter utilement les faits tokhariens à ceux des autres langues indo-européennes et, à plus forte raison, à ceux de l'indo-européen restitué lui-même. À vrai dire, suffisamment proches l'une de l'autre, les langues tokhariennes offrent un terrain de choix à l'application de la méthode comparative.

5.1

Il paraît naturel, à titre d'exemple, de conclure à l'existence d'un phonème proto-tokharien **w'* (**w* palatalisé : cf. §8.2.1) en se fondant sur la correspondance régulière A *w* = B *y*, devant voyelle antérieure (à l'exception de **i* bref) : cf. A *want*, B *yente* « vent » < **w'ente* (indo-européen **wēnto-*) ; A *malyw-*, B *mely-* « broyer, écraser » < **mel'w'(ä)y-* (indo-européen **mol[H]w-eye/o-*) ; A *wäs*, B *yasa* « or » < **w'äs-a* (indo-européen **[H_2]wes-H_2*), etc. De la même façon, on peut ramener à l'unité des formes à première vue divergentes comme A *klis-*, B *klänts-* « dormir » (**kläns-* < indo-européen **klm̥-s-*), A *es*, B *āntse* « épaule ») (**anse* < indo-européen **H_2emso-*) s'il est admis qu'une séquence tokharienne commune **-ns-* intérieure se développe en A *-ys-* (avec yodisation de *n*, puis monophtongaison), B *-nts* (naissance d'un *t* intercalaire). Du même coup s'éclaircit le rapport des désinences de génitif singulier A *-es* = B *-entse* (flexion thématique : **-ense*, cf. § 4.1.1) et A *-is* = B *-äntse* (flexion athématique : **-änse*).

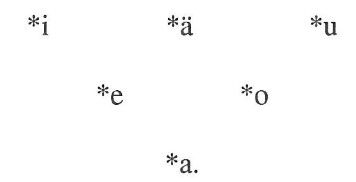
En morphologie encore, la comparaison des désinences d'accusatif pluriel A *-(a)s* = B *-(e)ṃ* à la flexion thématique ne nécessite aucun recours au témoignage d'autres langues indo-européennes pour reconstruire une finale **-(e)-ns* en proto-tokharien (indo-européen **-(o)-ns* : §4.1.1). Mais une différence telle que A *-eñc* ↔ B *-eṃ*, dans la marque de la troisième personne du pluriel actif (§ 4.2) apparaît irréductible : il faut compter, en tokharien commun, avec les reflets concurrents de l'indo-européen **-o-nti* (indicatif) et **-o-nt* (injonctif).

Dans l'ensemble, la langue B apparaît plus archaïque. Elle garde intactes un grand nombre de diphtongues anciennes, alors que la langue A

procède à une contraction généralisée ; et la fin de mot, gravement atteinte en tokharien A, y est beaucoup mieux préservée : cf. A *ekär* = B *aikare* « vide » < **aik(ä)re* (indo-européen **aigh-ro-*) ; A *top*, B *taupe* « mine » < **taupe* (indo-européen **dhoubo-*), etc.

5.2

Dans la reconstruction des voyelles proto-tokhariennes que nous proposons ci-dessous à titre d'illustration, nous avons choisi délibérément des symboles phonologiques proches des reflets de la langue B :



Les phonèmes du tokharien commun sont définis par les correspondances suivantes :

**i* (A *i* = B *i*) : A *wiki*, B *ikäṃ* « vingt » < **w'ikän(t)* (indo-européen **wīk'n̥t*) ;
**ä* (A *ä* = B *ä*) : AB *pär-* « porter » < **pär-* (indo-européen **bher-*) ;
**u* (A *u* = B *u*) : AB *ku* « chien » < **k(w)u* (indo-européen **k'wōn*) ;
**e* (A *a* = B *e*) : A *ak*, B *ek* « œil » < **ek(w)* (indo-européen **ok*^{w}*-*) ;
**o* (A *a* = B *o*) : A *pracar*, B *procer* « frère » < **procer* (indo-européen **bhrātēr*) ;
**a* (A *ā* = B *ā*) : A *āk*, B *āke* « bout, fin » < **ake* (indo-européen **ak'-os*).

5.3

À l'époque postérieure au tokharien commun, deux effets différents de l'accentuation donnent lieu, dans la langue B, à une nouvelle voyelle centrale *ə* (notée <a>), née de *ä* tonique d'une part, de *a* atone d'autre part (d'où les correspondances occasionnelles A *ä* = B *a* [*ə́*] et A *ā* = B *a* [*ə̀*] : cf. par exemple A *śäṃ* = B *śana* [*śə́nā*] « épouse » < **śäna* < indo-européen **g*^{w}*enH*$_2$: § 4.1.2, etc.).

D'autre part, les monophtongaisons survenues dans la langue orientale y créent les nouvelles voyelles *e* (de **ai*, **ei*, etc.) et *o* (de **au*, **eu*, etc.), qui exercent une pression latérale sur les **e* et **o* hérités et les font coïncider en *ə*, voyelle centrale de même aperture, que

l'écriture note (comme en B) par le caractère <a> (d'où les correspondances A *a* = B *e* et A *a* = B *o*). La labialisation des voyelles centrales A *a* et *ä* constitue une autre source de divergence phonétique entre les deux langues tokhariennes : au contact d'une labiale ou d'une labio-vélaire, A *a* (= B *e* ou *o*) et *ä* (= B *a* ou *ä*) sont susceptibles de s'arrondir en *o* et *u* respectivement : cf. A *koṃ* (acc. sg.) «chien» < A **kwaṃ* [B *kweṃ*] < **kwen* (indo-européen **k'won-m̥*); A *ñom* «nom» < A **ñam* [B *ñem*] (indo-européen *[H_1]*nē*[H_3]*mn̥* : §4.1.2); A *pont-* «tout» < A **pant-* [B *pont-*] < **pont-* (indo-européen **pānt-*); A *poke* «bras» < **pakwe* [B acc. *pokai*] < **pok(w)ai* (indo-européen **bhāg'hu-*); A *yuk* «cheval» < A **yäkw* [B *yakwe*] < **yäkwe* (indo-européen **ek'wo-*); A *pukäl* «année» < A **päkwäl* [B *pikul*] < **p'äkwäl* (indo-européen **pekʷ-l̥* «[temps de la] maturité»); A *puk* «tout» < A **päkw*/**wäkw* < **w'äkwe* (indo-européen **wik'wo-*), etc.

Les effets de l'accentuation en tokharien A ne sont guère visibles. Mais on y décèle l'action d'une sorte de balancement vocalique, affectant le degré d'aperture de *a* et de *ā* à l'intérieur du mot après voyelle initiale *e, o, a* ou *ā* (cf. A **tākār* > *tākar* «ils furent», à côté de B *takāre* = *tākā́re*).

6

Les développements qui suivent visent à synthétiser les principales tendances phonétiques qui caractérisent les langues tokhariennes à l'intérieur de la famille indo-européenne. Les progrès réalisés dans ce domaine, plus de quatre-vingts ans après l'étude fondatrice de Sieg et de Siegling (§1), illustrent bien la nécessité méthodologique de l'interaction incessante entre l'investigation étymologique et la fixation des lois phonétiques.

7. Le vocalisme

Les caractères vocaliques utilisés par l'écriture brāhmī ne posent pas de problèmes d'interprétation, si ce n'est dans le cas des signes notant

<a>, <ā> et <ä>. Dans l'état actuel de nos connaissances, il convient d'y voir le reflet des voyelles centrales [a] (ouverte), [ə] (semi-ouverte) et [ɨ] (fermée); les transcriptions tokhariennes de mots empruntés au sanskrit et à l'iranien montrent en tout cas que l'opposition <a> : <ā> n'est pas de nature quantitative. Le fait le plus marquant dans l'évolution du vocalisme tokharien, depuis la période préhistorique, est précisément l'égalisation de la durée des longues et des brèves (isochronie) – à coté de la tendance nettement marquée à l'altération du timbre sous l'influence des voyelles environnantes (Umlaut ou métaphonie).

7.1. L'isochronie

Le vocalisme indo-européen était défini par les paramètres de localisation, de degré d'aperture et de durée. Si l'opposition de quantité a été neutralisée en tokharien dans un processus appelé par d'aucuns « _the great vowel shift_ » (qui n'interdit pas bien entendu l'allongement ou l'abrègement allophoniques, accompagnant par exemple l'accentuation), les anciennes voyelles brèves et longues continuent à se distinguer par la qualité (transphonologisation). Ne sont dès lors jamais confondues en proto-tokharien les continuations de :

indo-européen _*a_ et _*ā_ > tokh. comm. _*a_ et _*o_
indo-européen _*e_ et _*ē_ > tokh. comm. _*ϕ (ä)_[3] et _*e_[3]
indo-européen _*o_ et _*ō_ > tokh. comm. _*e_ et _*a_ (_*-u_ en finale)
indo-européen _*i_ et _*ī_ > tokh. comm. _*ϕ (ä)_[3] et _*i_[3]
indo-européen _*u_ et _*ū_ > tokh. comm. _*ϕ (ä)_[4] et _*u_.

7.1.1. Exemples

A _pāk_, B _pāke_ « part » < _pake_ (indo-européen _*bhago-_); A _ānt_, B _ānte_ « surface, front » < _*ante_ (indo-européen _*anto-_); AB _āp_ « eau, rivière » < _*ap_ (indo-européen _*ap-_);

A _waṣt_, B _ost_ « maison » < _*wost_ (indo-européen _*wāstu-_ : §4.1); A _śna-c_ (all. sg.), B _śano_ (acc. sg.) « épouse » < _*śäno_ (indo-européen _*g_^{w}_enā-m_ : § 4.1.2);

A _śaru_, B _śerwe_ « chasseur » < _*śerwe_ (indo-européen _*g_^{w}_ērwo-_ « pourvu d'une lance »);

3. Avec, le cas échéant, prothèse d'un yod ou palatalisation de la consonne précédente.

4. Avec, le cas échéant, développement d'un _w_ prothétique.

A _wak_, B _wek_ « voix » < _*wek(w)_ (indo-européen _*wok_^{w}_-_) ; A _tkaṃ_, B _keṃ_ « terre, sol » < _*tken-_ (indo-européen _dhg'hom-_) ;
A _ysār_, B _yasar_ « sang » < _*yäsar_ (indo-européen _*[H]es[H]ōr_) ; A _wu_ (masc.) « deux » < _*wu_ (indo-européen _*dwō_ : §8.1.2) ;
A _wät_, B _wate_ « deuxième » < _*wäte_ (_*dwi-to-_) ; A _wäs_, B _wase_ « poison » < _*wäse_ (indo-européen _*wiso-_) ; A _lyäm_, B _lyam_ « lac » < _*lyäm_ (indo-européen _*li-mn̥_) ;
A _wir_ « jeune, vigoureux » < _*wire_ (indo-européen _*wīro-_) ;
A _tpār_, B _tapre_ « haut » < _*täpre_ (indo-européen _*dhub-ro-_) ; A _rtär_, B _ratre_ « rouge » < _*rätre_ (indo-européen _*rudh-ro-_) ; B _akwatse_ « aigu, sévère » < _*akwätse_ (indo-européen _*ak'u-tyo-_) ;
A _nuṃ, nunak_ « de nouveau » < _*nune(-k)_ (indo-européen _*nū-no-_), etc.

7.1.2

Le seul contexte où les anciennes longues et brèves puissent coïncider est fourni par le traitement des diphtongues proto-tokhariennes : sous l'accent, _*ei_ aboutit, tout comme _*ēi_, à tokh. comm. _*ey_ (A _e_, B _ai_) et _*eu_ se fond pareillement dans _*ēu_ pour donner tokh. comm. _*ew_ (A _o_, B _au_) ; cf. A _tre_, B _trai_ « trois » < _*trey_ (et non †_tri_ < †_träy_) (indo-européen _*tréy-es_) ; A _klyos-_, B _klyaus-_ (prés.) « écouter, entendre » < _*klyews-_ (et non †_klyus-_ < †_klyäws-_) (indo-européen _*k'léu-se/o-_), etc.

7.1.3

La voyelle centrale fermée _ä_ n'a pas à proprement parler de réalité phonologique ; c'est une voyelle d'anaptyxe, qui se développe automatiquement à l'intérieur de groupes consonantiques complexes. Elle s'amuït normalement en syllabe ouverte ; très sensible à l'influence des sons environnants, elle prend les timbres respectifs _i_ en contexte palatal et _u_ en contexte vélaire : cf. A _kuryar_ = B _karyor_ « commerce » < _*k(w)riyor/*k(w)räyor_ (indo-européen _*k_^{w}_riyār_) et A °_klyu_ = B °_kälywe_ « gloire » < _*kl'uwe/*kl'äwe_ (indo-européen _*k'lew-os_). En dehors du traitement de indo-européen _*e_, _*i_ et _*u_, on trouve, comme autre source de _*ä_ en tokharien, la résolution des liquides et nasales voyelles _*-r̥-_, _*-l̥-_, _*-m̥-_, et _*-n̥-_ : cf. A _śtärt_, B _śtart(t)e_ « quatrième » < _*śät(w)ärte_ (indo-européen _*k_^{w}_et(w)r̥-to-_ ; A _känt_, B _kante_ « cent » < _*känte_ (indo-européen _*k'm̥to-_) ; B _walkwe_ « loup » < _*wälkwe_ (indo-européen _*wl̥k_^{w}_o-_), etc. Mais la séquence _*-äR-_ du « tokharien » commun remonte peut-être elle-même à un traitement plus ancien _*-uR-_, comme le suggèrent des cas comme B _walke_ « longtemps » < _*wälke_ < _*dwälke_ < _*dulke_, avec chute régulière de _*d_ devant _*w_ (§8.1.2) (indo-européen _*dl̥_H_1_-gho-_) ; A _wär_, B _war_ « eau » < _*wär_ < _*wädwär_ < _*udur_ (indo-européen _*ud-r̥_), etc.

7.2. La métaphonie

Dès l'époque du tokharien commun, beaucoup de phonèmes vocaliques ont subi l'action assimilatrice de voyelles subséquentes ; par la suite, les phénomènes d'umlaut se sont poursuivis au point d'obscurcir nombre de rapports lexicaux entre les langues A et B. Dans ce qui suit, nous voudrions (sans viser à l'exhaustivité dans un domaine aussi complexe) illustrer la tendance tokharienne à la métaphonie par quelques cas typiques :

7.2.1

L'Umlaut de **u* sur **e* (d'où tokharien commun **ö*) : cf. A *okät*, B *okt* « huit » < **ökt(ä)* < **ekt(u)* (indo-européen **ok'tō*) ; AB *or* « bois » < **ör(ä)* < **er(u)* (indo-européen **doru-*, avec chute de la dentale initiale) ; AB *yok* « couleur, cheveu » < **yök(ä)* < **yek(u)* (indo-européen **yēg*^{w}*u-*) ; B *ṣñor* « tendon » < **ṣñö(w)ur* < **ṣñe(w)ur* (indo-européen **snē-wr̥*) ; B *mot* « alcool » < **möt(ä)* < **met(u)* (indo-européen **mēdhu-*, en alternance apophonique avec le thème **medhu-*), etc. La correspondance A *o* = B *o* définit donc, pour le proto-tokharien, un phonème **ö* distinct de **o* (A *a* = B *o* : §5.2)[5].

7.2.2

L'Umlaut de **a* sur **e* (d'où tokharien commun **a*) : cf. B *śtwāra* (fém. de *śtwer* [= A *śtwar*] « quatre » < **śtwara* < **śtwera* (indo-européen **k*^{w}*etwor-*H_2) ; B *āsta* (plur.) « os » < **asta* < **esta* (indo-européen **(H)ost-*H_2) ; B *ārwa* (plur. de *or* « bois ») < **arwa* < **erwa* (indo-européen **doru-*H_2) ; B *yākwa* (plur. de *yok* « couleur, cheveu ») < **yakwa* < **yekwa* (indo-européen **yēg*^{w}*u-*H_2) ; B *ṣñaura* (plur. de *ṣñor* « tendon ») < **ṣñawra* < **ṣñewra* (indo-européen **snē-wr̥-*H_2), etc.

7.2.3

L'Umlaut de **o* sur **e* (d'où tokharien commun **o*) : cf. A *klawa-*, B *klowo-* « annoncer » < **klowo-* < **klewo-* (indo-européen **k'low-ā-* « faire entendre ») ; A *praska(-ñi)* « craintif », B *prosko* « crainte » < **prosko* < **presko* (indo-européen **prok-sk'-ā*) ; B *somo* (fém. de *ṣe*) « une seule » < **somo* < **semo* (indo-européen **som-ā*), etc.

5. Il est possible qu'il faille reconstruire parallèlement pour la série antérieure une voyelle **ë*, née de **e* devant **i*, ce qui rendrait compte de l'accord A *e* = B *e* (cf. AB *yneś* « manifestement » < **y(ä)nëś(ä)* < **y(ä)neś(i)* < indo-européen **en-ok*^{w}*i* à côté de la désinence dans A *ākeñc* [B *ākeṃ*] « ils conduisent » < **-ëñc(ä)* < **-eñc(i)* [indo-européen **ag'-onti*], déjà étudiée au §5.1).

8. Le consonantisme

Les signes consonantiques que l'écriture brāhmī utilise pour noter les mots tokhariens hérités sont :

<p>, <t>, <k> (occlusives),
<c> (palatale) et <ts> (affriquée),
<ś>, <ṣ>, <s> (sifflantes),
<r>, <l>, <ly> (liquides),
<m>, <n>, <ñ>, <ṅ> (nasales) et
<w>, <y> (semi-voyelles),

dont l'interprétation phonétique ne rencontre pas de problème particulier. Par rapport au système indo-européen, le consonantisme tokharien se signale avant tout par la perte des traits de sonorité et d'aspiration d'un côté et par l'introduction de la palatalisation comme marque phonologique ; par ailleurs, une dentale affriquée _ts_ s'est développée dans des conditions en partie non élucidées.

8.1. La mutation consonantique

La principale caractéristique du consonantisme tokharien comparé à celui de l'indo-européen est sans conteste la neutralisation de l'opposition de sonorité ; à cette réduction s'ajoute, pour ce qui est des occlusives proprement dites, la perte de l'aspiration. Pour illustrer ce vaste processus de mutation articulatoire, on se contentera ici de quelques étymologies seulement : A _käntu_, B _kantwo_ « langue » < _*käntwo_ (indo-européen _*dn̥ǵhwā_, avec métathèse des occlusives) ; AB _pär-_ « porter » (indo-européen _*bher-_ : §5.2) ; B _twere_ « porte » (indo-européen _*dhworo-_) ; AB _āk-_ < _*ak-_ (indo-européen _*ag'-_ : §4.2) ; A _kär_, B _käre_ « dignité, valeur » _*käre_ (indo-européen _*g_^{w}_r̥_H_2_-os_ « poids »), etc.

8.1.1

Rien n'exclut à vrai dire que les phonèmes sourds comme /p/, /t/, /k/ aient connu des variantes contextuelles sonores et/ou constrictives (par exemple en position intervocalique ou au contact d'un liquide) ; un reflet écrit de cette allophonie se trouve en tout cas dans le passage fréquent de la labiale _p_ à _w_ entre voyelles ainsi qu'après _r_ et _l_. On peut supposer également que le redoublement de certaines consonnes (intervocaliques ou devant _r_) sauvegardait la prononciation sourde et qu'ainsi l'opposition

géminée : simple pouvait se changer en une opposition sourde : voisée, ce qui réintroduisait dans la langue une sonorité phonologique.

8.1.2

Quelques indices montrent que, dans l'histoire pré-tokharienne, les oppositions indo-européennes de sonorité et d'aspiration avaient toujours cours. A regarder de près les traitements tokhariens des dentales, on constate que seule la sonore **d* s'assimile ou disparaît devant semi-voyelle, liquide ou nasale, alors que **t* et **dh* se maintiennent dans le même contexte : A *pe*, B *paiyye* « pied » < **peyye* ou **payye* (indo-européen **pō̆d-yo-*) ; A *swār*, B *swāre* « doux, agréable » < **sware* (indo-européen *suH_2d-ro-*) ; A *wu* (masc.), *we* (fém.), B *wi* (masc.-fém.) « deux » (indo-européen **dwō*, **dwai*, **dwoi*), etc. De façon comparable, *dh* patalisé aboutit à *ts*, c'est-à-dire à un produit différent de la patalisation de **t* et **d* : B *tsik-* « façonner, former » (indo-européen **dheigʹh-*) ; A *tsopats* « grand » < **tsaupets(ts)e* (indo-européen **dhéubo-tyo-*). En ce qui concerne les labiales, à la différence absolue de **p* et de **b*, l'aspirée **bh* s'amuit derrière **m* : A *kam*, B *keme* « dent » (indo-européen **g'ombho-*) ; B *sáma* « se trouva » < **śäma* (indo-européen **stembhH-t*), etc.

8.2. La palatalisation

Comme dans nombre d'autre langues indo-européennes, les consonnes vélaires et dentales ont subi en tokharien une mouillure devant voyelle antérieure, en l'occurrence **ē*, **e*, **ī̆* et (seulement pour les dentales) **i*[6]. Les nouvelles articulations **k'*, **t'*, **s'*, **n'* et **l'* ont évolué vers *ś, c, ṣ, ñ* et *ly*, qui ont accédé au statut de phonèmes dans la mesure où, dans la suite, la voyelle palatalisante a disparu ou s'est modifiée : ainsi A *śpāl* « tête », adj. dér. AB *śpāl-u* « excellent » < **k'äpal* (indo-européen **ghebhH-ōl*) ; A *pśär* « aigu » < **p'äk'äre* (indo-européen **pik'-ero-*) ; B *cake* « rivière » < **t'äke* (indo-européen **tek-os* « courant ») ; A **śre*, B *ścirye* « étoile » < **s't'är-ye* (indo-européen *[*H*]*ster-yo-*) ; A *ṣäk*, B *ṣkas* « six » < **s'äks* (indo-européen **sek's*) ; A *ñu*, B *ñuwe* « nouveau » < **n'äwe* (indo-européen **newo-*) ; A *ñkät*, B *ñakte* « dieu » < **n'äk(w)te* (indo-européen **k'wento-* « saint », avec interversion des consonnes radicales) ; B *kälyp-* « voler » < **k(ä)l'äp-* (indo-européen **klep-*), etc.

6. Les mêmes consonnes ont été également altérées devant la semi-voyelle **y* ; le résultat du groupe **Cy* donne en général une palatale géminée du type *śś* < **ky*, *ṣṣ* < **sy*, *ññ* < **ny*, *lyly* < **ly*, etc. Pour le groupe **ty* > *ts*, cf. §8.3.

8.2.1

Même la prononciation des labiales a dû, dans le même contexte, être affectée de mouillure; mais ces consonnes ont été généralement dépalatalisées. En koutchéen, toutefois, l'articulation du type **p'* et **m'* devant *ä* a laissé quelques traces en modifiant cette voyelle en *i* : B *pīle* [A *päl*] «blessure» < **p'äle* (indo-européen *[H_2]*pel-os*); B *piṅkte* [A *pänte*] «cinquième» < **p'äṅkte* (indo-européen **penk*^{w}*-to-*); B *mit* «miel» < **m'ät* (indo-européen **medhu-*), etc. D'autre part, en tokharien A, la semi-voyelle **w'* a été dépalatalisée, alors que la langue B laisse se poursuivre le processus de **w'* vers *y* : A *wkäṃ*, B *yakne* «manière, façon» < **w'äkne* (indo-européen **weg'h-no-* «conduite»); A *walyi*, B *yelyi* (plur.) «vers» < **w'el'äy* (indo-européen **wēl-eyes* ou **wēl-iyes*), etc.

8.2.2

A l'initiale absolue, les voyelles antérieures développent un yod prothétique, issu d'un 'on-glide' palatal : A *yuk*, B *yakwe* «cheval» < **yäkwe* (indo-européen **ek'wo-* : cf. §5.3); A *yärk*, B *yarke* «hommage» < **yärke* (indo-européen **erk-os*); A *ytār*, B *ytārye* «chemin» < **yätar-* (indo-européen *[H_1]*i-tōr*), etc.

8.3. La genèse de l'affriquée *ts*

La dentale *ts* est, avec la voyelle *ä*, un des éléments les plus typiques du phonétisme tokharien. Son origine reste controversée, sauf dans les cas où *ts* résulte d'un groupe **ty* (cf. A *pats*, B *petso* [acc. sg.] «époux» < **petso* < indo-européen **pot-yā-m*) ou de **dh* palatalisé devant **ē̆* ou **ī̆* (cf. §8.1.2). D'autres exemples laissent entrevoir que *ts* est également le traitement régulier de la sonore **d* devant voyelle, à côté du produit palatalisé *ś* :

**d*-V- > tokharien commun **ts-*
**d'*-V- > tokharien commun **ś-*

(cf. A *tsäk-* «tirer dehors» < **däk-* < indo-européen *duk-*; B *tsāk-* «mordre» < **dak-* < indo-européen **dak'-*; A *śäk*, B *śak* «dix» < **d'äk* < **dek'-m̥*, etc.).

Bibliographie

ADAMS D.Q. (1988). – *Tocharian Historical Phonology and Morphology*, New Haven.

KRAUSE W. et THOMAS W. (1960-1964). – *Tocharisches Elementarbuch*, I. *Grammatik.* – II. *Texte und Glossar*, Heidelberg. [description complète des deux langues tokhariennes, suivie d'un dictionnaire].

PINAULT G.-J. (1989). – «Introduction au tokharien» (*LALIES.* Actes des sessions de linguistiques et de littérature, 7), Paris, p. 1-224. [remarquable synthèse, abordant successivement les données externes, la phonologie, la morphologie nominale et verbale, le commentaire de textes].

SIEG E. et SIEGLING W. (1908). – «Tocharisch, die Sprache der Indoskythen. Vorläufige Bemerkungen über eine bisher unbekannte indogermanische Literatursprache», *Sitzungsberichte der Königlichen Preussischen Akademie der Wissenschaften*, Berlin, p. 915-932. [étude «inaugurale»].

THOMAS W. (1985). — *Die Erforschung des Tocharischen* (1960-1984). – Stuttgart. [bibliographie analytique et critique de l'ensemble des travaux tokharologiques modernes].

Tocharian and Indo-European Studies (éd. J. Hilmarsson), Reykjavík (1987 et suiv.). – [la seule revue de tokharologie, doublée d'une «Supplementary Series» de monographies].

VAN WINDEKENS A.J. (1976). – *Le tokharien confronté avec les autres langues indo-européennes*, I. *La phonétique et le vocabulaire*, Louvain. [les p. 139-646 contiennent un lexique étymologique].

WINTER W. (1984). – *Studia Tocharica. Selected Writings*, Poznań. [important recueil d'articles].

CHAPITRE 5

L'INDO-IRANIEN

Manfred MAYRHOFER

1

Parmi les grands groupes de dialectes de l'indo-européen, le groupe indo-iranien occupe une place à part. Plus qu'en tout autre cas, le composé « indo-iranien » renvoie à une *unité* originelle : les deux groupes de langues de l'indien ou mieux de l'indo-aryen (ci-dessous 3.1.3) et de l'iranien (ci-dessous 3.3), dans leurs fixations les plus anciennes, la langue indienne archaïque du Rig-Veda et la langue iranienne archaïque de l'Avesta sont encore si semblables que des tranches entières de textes conservés peuvent être « transposées » d'une langue dans l'autre, moyennant l'application d'une petit nombre de lois phonétiques (ci-dessous 1.1.1-1.1.3) ; ainsi, par tous leurs éléments linguistiques (vocabulaire, grammaire, syntaxe), ils ramènent à une « protolangue indo-iranienne » unique. Quand une proximité dialectale est telle qu'on ne peut l'expliquer que par une unité précédente, cela ne saurait étonner dans le cas de parlers géographiquement voisins, tardivement séparés ; au contraire, ce qui est si fascinant dans le cas

de l'indo-iranien, c'est que ces dialectes d'une seule protolangue indo-iranienne appartiennent à des pays et à des cultures différents dont les destins historiques sont totalement séparés depuis bien des siècles : ici le plus ancien livre sacré de l'hindouisme indien; là les hymnes du réformateur iranien Zarathustra ou bien les incriptions des rois des rois perses comme Darius I[er] et Xerxès ! Un explorateur contemporain de textes difficiles de l'Avesta m'a un jour avoué traduire les textes *iraniens* anciens en compulsant un dictionnaire du Veda, *indien* ancien. Cette anecdote jette une lumière sur la situation particulière des textes indo-iraniens : supposons qu'on parvienne (mais il n'en est pas question) à analyser la forme actuelle du roman en Dacie, le roumain, à l'aide d'un lexique et d'une grammaire du roman ibérique, de l'espagnol ou du portugais.

1.1.

Je tenterai de démontrer qu'il est possible de transposer des phrases entières d'une langue indo-iranienne ancienne dans une autre par un unique exemple; il marquera suffisamment par quoi l'iranien le plus ancien diffère du plus ancien indo-aryen (1.1.2), à quel point cette reconstruction était encore proche de l'indo-iranien (1.1.3), et en quoi cet indo-iranien se différencie de l'indo-européen (1.1.4.1-2).

1.1.1.

Dans l'hymne à Mithra de l'Avesta (Yašt 10.6) il est écrit :

1.1 *təm amauuaṇtəm yazatəm*
1.2 *sūrəm dāmōhu səuuištəm*
1.3 *miθrəm yazāi zaoθrābiiō*

Traduction :
1.1 Ce puissant dieu
1.2 Le fort, parmi les vivants le plus fort,
1.3 Mithra, je veux honorer par libations.

1.1.2.

Chaque mot, chaque cas, chaque terminaison verbale, chaque détail de l'ordre des mots est transposable dans la langue du Rig-Veda. La phrase ci-dessous aurait été totalement compréhensible pour un Indien de l'époque védique :

(1) *tám ámavantam yajatám*
(2) *śū́ram dhā́masu śáviṣṭham*
(3) *mitrám yajai hótrābhyaḥ*

1.1.3.

Si l'on reconstruit, à partir de là, un texte « indo-iranien », toutes les valeurs phonétiques sont compréhensible à partir d'une transcription du védique, à l'exception toutefois de **ć* = [*tś*] et **ǰ* = [*dź*] archétypes d'av. *s* = véd. *ś* = nuristani *ts* (cf. 2.1.2.2), et, respectivement, d'av. *z* = ved. *j* = nuristani *dz*. Ces trois lignes se présentent alors comme suit :

(1) **tám *ámavantam *yaǰatám*
(2) **ćū́ram *dhā́masu *ćávištham*
(3) **mitrám *yaǰā́i *ǰháutrābhyas*

1.1.3.1.

Cette petite expérience montre clairement que la forme phonétique de l'indo-iranien diffère peu de celle du védique. Quelques innovations de l'iranien sont visibles, ainsi *h<*s* (ligne 2 *dā́mōh̠u* = *dhā́mas̠u* ; cf. av. *hapta* « sept » = védique *saptá*, latin *septem* etc.), *d*, *b* de **dh*, **bh* (à nouveau : *d̠āmōhu* = *d̠h̠ā́masu* ; ligne 3 *-b̠iiō* = véd. *-b̠h̠yaḥ*), *θr* de **tr* : *miθ̠ra-* = *mitrá-*). Dans av. *zaoθrā-* « offrande, libation » = véd. *hótrā-*, il semble que **ǰ* (av. *yazāi*) et **ǰh* se sont confondus en iranien, de même que **d* et **dh*, **b* et **bh* ; mais le **ǰ* a disparu devant *-h-* en indo-aryen ancien, tandis que l'iranien n'a perdu que l'aspiration (véd. *h-* = av. *z-*), et la diphtongue *-ao-* de l'avestique est plus proche de l'indo-iranien (**ǰháutrā-*) que ne l'est l'*o* monophtongue du védique (*hótrā-*) ; cela vaut également pour la diphtongue longue iranienne dans *yazā̠i̠*, qui est plus archaïque que la diphtongue normale du védique dans *yaja̠i̠*. Ainsi, comme il fallait s'y attendre, ce n'est pas toujours l'indo-aryen, mais parfois l'iranien qui est plus proche du niveau antérieur commun (voir 3.4). Des divergences moindres affectent les finales (ainsi avest. *-əm* = véd. *-am* ; avest. *-ō*, véd. *-aḥ* ; les cas d'Umlaut (*dāmō̠hu̠* pour **dāmahu* = véd. *dhā́ma̠su̠*) ; ou des modifications minimes du vocalisme dans un environnement donné (ligne 2, avest. *səuuištəm* pour **sauuištəm*= véd. *śáviṣṭham*). A l'occasion, la différence entre ces langues étroitement apparentées n'est que graphique ; ainsi l'écriture littérale avestique note *amauuaṇt-* « puissant » pour une séquence phonétique dont la phonation était identique à celle de ce que l'indien écrit *amavant-*. (Des iranistes du passé écrivent même en avestique *amavant-* ; mais ce n'est pas admissible en philologie, car les lettres *uu* et *ṇ* - qui, dans l'écriture avestique, diffèrent de *v* et *n* !- doivent naturellement demeurer visibles dans une transcription scientifique).

1.1.3.2.

Ce qui importe plus que ce petit nombre de divergences, c'est bien entendu la grande proximité des plus anciennes langues de culture de l'Inde et de l'Iran qui ressort de cet exemple. Ces coïncidences remontent à une

protolangue commune dont ces deux grandes langues littéraires ne sont que des dialectes si étroitement apparentés que leurs locuteurs se seraient sans doute compris entre eux dans la plupart des cas. Et ce bref texte de l'Avesta, avec ses correspondants complètement conservés dans la langue du Rig-Veda indien, nous apporte un autre enseignement important. Dès ces trois lignes, nous trouvons des mots pour « honorer par un culte » (*yazāi* = *yajai*) et « digne d'adoration, dieu » (*yazata-* = *yajatá-*), pour « libation, offrande » (*zaoθrā-* = *hótrā-*), et pour le nom d'un dieu (*miθra-*= *mitrá-*); cela suggère que le vocabulaire indo-iranien reflète un *monde commun de la religion et du sacrifice* : des syntagmes se correspondent dans la langue de la prière (par exemple avest. *ustānazastō...nəmaŋhā* « tenant les mains levées écartés en signe d'adoration » = véd. *uttānáhasto námasā*, indo-iranien **uttānáȷ́hastas *námasā*, geste d'orant déjà indo-iranien) ; et des noms des plusieurs dieux, voire de généalogies divines entières coïncident (par exemple avest. *yima-*, fils de *vīuuahuuaṇt-* nom d'un roi mythique des origines = véd. *yamá-*, fils de *vivásvant-*, nom du roi du royaume des Morts ; il existe donc un mythe indo-iranien de **yamá-* fils de **vivásvant-*).

1.1.4.

Cette forte proximité des langues indo-aryenne et iranienne, cette communauté des univers religieux et culturels de l'indo-iranien a pour pendant un niveau indo-européen préalable nettement distinct de l'indo-iranien ; les innovations linguistiques communes par rapport à l'indo-européen ont contribué à dégager l'unité linguistique du niveau précédant immédiatement l'indo-aryen et l'iranien, et, dans une période assez longue de vie commune, avant la séparation des peuples, la religion et la culture originairement communes. Revenons une dernière fois à notre petit exemple (1.1.1-3) :

1.1.4.1.

Pour quelques éléments du texte indo-iranien, il y a des précédents indo-européens sûrs. Ainsi le premier mot **tám* correspond à grec τόν et suppose un accusatif singulier masculin indo-européen **to-m* « ce, cet » ; **dhā́masu* (ligne 2) est le locatif pluriel de **dhā́-man-* « créature » et on lui compare pour le degré vocalique et le suffixe gr. ἀνά-θη-μα « ce qui est placé, offrande ». Là apparaît l'innovation typique du vocalisme indo-iranien : à la multiplicité vocalique, héritée de l'indo-européen, que présentent le grec et le latin *e/o/a* et *ē/ō/ā*, répondent en indo-iranien *a* et *ā* par un processus radical de « déphonologisation » ; à l'appui des exemples qu'apporte notre texte (pour *a* = grec *o* ; pour *ā* = grec *ē*) on ajoutera véd. *ásti* « il est » = latin *est* ; véd. *ájati* « il pousse » = lat. *agit* ; véd. *dā́nam* « don » = lat. *dōnum* ; véd. *bhrā́tar-* « frère » = lat. *frāter*. Ce processus d'évolution phonétique entraîne une prédominance, dans les

langues indo-iraniennes, du timbre *-a-/-ā-*, tel qu'il apparaît dans notre texte (1.1.3); il caractérise aussi les nom tirés des anciennes cultures indienne et iranienne, par exemple : familiers à tout individu cultivé, ceux des grandes épopées indiennes *Mahābhārata* et *Rāmāyaṇa*; le nom de famille du Bouddha : *Gautama*, et même le surnom d'un homme politique de notre siècle, en sanskrit *Mahātmā* (Gandhi); ou bien le nom du prophète iranien ancien *Zaraθuštra* («Zoroastre») et des rois des rois perses *Dārayavauš* («Darius») et *Xšayāršā* (Xerxès).

1.1.4.2.

Notre texte montre une autre innovation indo-iranienne par rapport à l'indo-européen : **ćū́ra-*«fort», **ćávištha-* «très fort» (ligne 2) remontent à indo-européen **k̑uH-ro, *k̑euH-isto-*, cf. grec κῦριος «maître, souverain»; **yaj́-* «honorer par un culte» (1.1.3.1) remonte à indo-européen **Hiag̑-* , cf. grec ἅγιος «saint»; et **j́háutrā-* «libation» (1.1.3.1) à indo-européen **g̑heu-* «verser», cf. gr. χεῦ-μα «ce qui est versé, courant», got. *giutan* «verser», all. *giessen*, etc. Des phonèmes palataux indo-européens **k̑*, **g̑*, **g̑h* représentés par *k, g*, etc., dans les langues «centum» comme le grec, le latin, le germanique etc., procèdent dans la langue «satəm» qu'est l'indo-iranien les sibilantes **ć*, **j́*, **j́h* (v. ci-dessus 1.1.3).

La forte modification du vocalisme (*ĕ, ŏ, ă* > *ă*, ci-dessous 1.1.4.1) et l'apparition de sibilantes procédant d'occlusives indo-européennes sont les innovations phonétiques communes les plus frappantes de l'indo-aryen et de l'iranien; du reste, ces deux langues sont des dialectes indo-européens d'une haute antiquité et précocement attestés qui jouent un rôle important dans la reconstruction de l'indo-européen.

2

La dénomination d'«indo-iranien», appliquée au niveau commun original de ces deux langues anciennes attestées de bonne heure, semble donc exacte et suffisante. On verra en un autre lieu de ce livre (p. 73) que pour «indo-iranien», on peut employer et emploie l'expression «aryen», d'après une auto-dénomination des deux peuples indo-iraniens (véd. *ā́rya-* «Indien appartenant à l'une des trois grandes castes supérieures»; vieux perse *ariya-* «iranien», persan mod. *Īrān* < **ariyānām* «pays des Iraniens»). La clarté de l'expression «indo-iranien» est peut-être remise en cause par l'existence d'un petit groupe d'idiomes tardivement attestés, les

langues du Nouristan (ci-dessous 2.1-2.2) qui montrent des affinités tant avec l'indo-aryen qu'avec l'iranien. Parmi les essais tentés pour interpréter ce phénomène (ci-dessous 2.2) se situe une conception à laquelle nous renvoyons dès maintenant : les langues du Nouristan continuent un troisième dialecte de l'indo-iranien, lequel prend place à côté de l'indo-aryen et de l'iranien; les partisans de cette opinion proposent de nommer ce groupe de langues : indo-irano-nouristanais.

2.1.

On appelle *Nūristānī* (langues du Nouristan) cinq langues de petites tribus de l'Hindou-kouch : *Aškun*, *Kati*, *Prasun*, *Waigali* et, proche du *Waigali*, la « langue des trois villages » *Tregami*. Dans les ouvrages spécialisés anciens (et à l'occasion, chez des auteurs d'aujourd'hui), on trouve pour « langues du Nouristan » l'expression « langues Kafir » (*Kāfirī*) qu'on voudrait voir éviter pour une raison de fait : *kāfir* (« mécréant, païen, non musulman ») était donné aux habitants du Nouristan avant leur conversion à l'Islam; comme ils sont devenus musulmans depuis la fin du XIX[e] siècle, le nom est devenu inadéquat et fut remplacé (Nūristān, « pays de l'illumination » pour *Kāfiristān* « pays païen »). Notre connaissance de ces langues du Nouristan ne remonte qu'au siècle dernier et surtout aux travaux de pionniers intervenus dans notre siècle; si l'on considère l'importance des renseignements que ces langues semblent apporter pour l'indo-iranien, cette attestation très tardive est fort regrettable (2.1.2.2.1, 2.2)

2.1.1.

D'une part, bien des éléments des langues du Nouristan, notamment dans le vocabulaire et surtout dans le vocabulaire religieux, donnent une impression d'un *indo-aryen* très archaïque. Quand on reconstruit les noms de dieux et de démons qui se sont conservés dans les langues de ces hommes si tard venus à l'Islam, on se trouve reporté trois mille ans en arrière, à l'époque du Rig-Veda, voire à une époque antérieure. Le dieu suprême a, dans les langues du Nouristan, des noms comme : aškun *imrā́*, prasun *yumrā́* (etc.), qui renvoient à *yama-rāja-* « Yama (ci-dessus 1.1.3.2) le roi »; *yama-rā́jya-* « l'empire de Yama » apparaît dans l'Atharva-Veda. Le serviteur du dieu suprême *yumrā́* s'appelle en Prasun *giwē̍ṣ*, à quoi correspond en Waigali *giwiš* « diable »; ces noms continuent l'adjectif du Rig-Veda *gav-iṣá-* « avide de bœufs ». Un autre dieu s'appelle en Prasun *indr*, c'est le dieu *índra-*, si souvent glorifié dans le Rig-Veda, qui apparaît dans le traité du Mittani (ci-dessous 3.2.2); à **indra-iṣṭi-* « coup d'Indra » remonte le kati *indríç* « tremblement de terre ». Je citerai deux noms parmi les démons bons ou mauvais : kati *dewutr* « fée » d'indo-aryen *deva-putrī-* « fille de Dieu », et kati *yuṣ* « démone mauvaise », cf. véd. *yakṣá-*

« apparition, fantôme, revenant ». Ce mince échantillon semble confirmer ce que dit Gérard Fussman : les peuples du Nouristan « sont les descendants d'Indiens qui n'auraient jamais pénétré dans le Panjab, ni dans la vallée du Gange, autrement dit : *des populations de langue presque indo-aryenne n'ayant subi ni l'influence de l'Inde non aryenne, ni celle de la théologie brahmanique* » (souligné par Fussman).

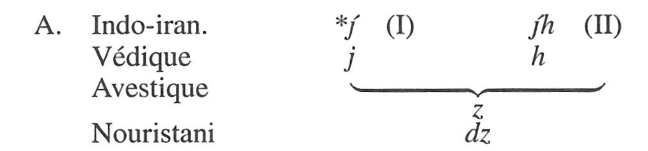
2.1.2.

D'autre part, les langues du Nouristan présentent aussi des analogies avec *l'iranien.*

Notre exemple (ci-dessus 1.1.3.1) nous a enseigné que l'indo-iranien **ȷ́* et **ȷ́h* (indo-européen **ĝ* et **ĝh*) coïncident en iranien (cf. *yaz-* et *zaoθrā-*) tandis qu'en indo-aryen, ces deux valeurs restent séparées (véd. *yaj-*, mais *hótrā-*). Il faut de surcroît constater ici que l'indo-iranien **ǰ* et **ǰh* (= indo-européen *$g^{(w)}$ et *$g^{(w)}h$ devant *e* ou *i* subséquent) coïncident en iranien, mais en un *ǰ* différent de **ȷ́*, **ȷ́h* (avest. *z*) ; en indo-aryen, les deux valeurs restent séparées mais ne diffèrent pas des résultats de **ȷ́*, **ȷ́h* ; nous avons en védique *j* pour **ǰ* (et **ȷ́*), *h* pour **ǰh* (et **ȷ́h*). Exemples : indo-européen *g^{w}*i̯éh*$_2$- « corde d'arc » (cf. grec βιός « arc »), indo-iranien **ǰiyáH-*, donne av. *ǰiiā-*, véd. *jiyā́-* ; indo-européen *g^{w}*hen-* « frapper, tuer » (hitt. *kuen-* « tuer »), indo-iran. **ǰhan-*, donne av. *ǰan-*, véd. *han-*.

2.1.2.1.

Les langues du Nouristan se comportent comme les langues iraniennes *dans cette évolution phonétique certainement très ancienne*, non comme les langues indo-aryennes. Je tenterai de clarifier cela sous la forme de deux tableaux.



(A. I. : **ȷ́ā́nu-* « genou », indo-européen **ǵónu-*, grec γόνυ, véd. *jā́nu-*/av. *zānu-*, kati *dzå*, transcription *jå* ;
A. II. : **ȷ́him-* « froid, neige, hiver », indo-européen **ǵhim-/*ǵheim-*, grec χειμών, russ. *zimá* « hiver » ~ ved. *himá-* « froid », cf. « Himālaya »/ av. *zəm-* « hiver » kati *dzīm*, transcription *jīm* « neige ».

B.	Indo-iran.	*ǰ (I)	*ǰh (II)
	Védique	*j*	*h*
	Avestique	ǰ	
	Nouristani	ǰ	

(B.I. : **ǰiyáH-* (ci-dessus 2.1.2) ~ véd. *jiyā́-*/ avest. *ǰiiā-*, kati *ǰī* «corde d'arc»;
B. II. : **ǰhan-* (ci-dessus 2.1.2) ~ véd. *han-*/ avest. *ǰan-*, kati *jäř-* «tuer»).

L'iranien et le Nouristani distinguent donc indo-iran. **ȷ́(h)* et **ǰ(h)* (*z/dz* ~ *ǰ/ǰ*); l'indo-aryen fait coïncider **ȷ́* et **ǰ* (véd. *j*), et de même **ȷ́h* et **ǰh* (véd. *h*), mais en revanche **ȷ́/*ǰ* de **ȷ́h/*ǰh* (*j* ~ *h*), tandis que l'iranien et le Nouristani ne font pas de différence entre aspirée et non aspirée. Il est difficile de ne pas considérer cette coïncidence des langues du Nouristan avec l'iranien, à la différence de la stricte distinction du védique, comme un processus évolutif très ancien.

2.1.2.2.

Il faut mentionner, en outre, l'évolution des palatales (**ȷ́*, **ȷ́h*) passant à *dz* en Nouristani – et de même dans le cas de la sourde, indo-iranien **ć* (ci-dessus 1.1.3) qui donne nour. *ts*, cf. indo-européen **déḱm̥* «dix» (grec δέκα, latin *decem*), indo-iranien **dáća* (véd. *dáśa*, avest. *dasa*) = kati *duts* (transcr. *duc*) : il semble exister un accord de plus avec l'iranien. Étant donné qu'indo-iranien **ć* donne *s* dans une partie des dialectes iraniens (par exemple en avestique), **θ* dans une autre partie (**daθa* > persan mod. *dah* «dix»), qu'en outre **ȷ́(h)* > *z* dans une partie des dialectes (ainsi l'avestique), **d* dans une autre (avest. *zəm-* [ci-dessus 2.1.2.1], kumzārī (Iran de Sud-Est) *dimestān* «hiver»), on postule iran. commun **ts*, **dz* pour *s/θ, z/d*; c'est là une évolution qui semble coïncider avec celle du Nouristani : indo-iranien **ć*, **ȷ́(h)* > nour. *ts*, *dz* (*contre* véd. *ś, j/h*).

2.1.2.2.1. Tout cela n'est avancé qu'avec une extrême réserve à laquelle nous contraint le fait que les langues du Nouristan sont attestées plus de trois mille ans (!) après les plus anciennes langues indo-iraniennes; dans ces «millénaires obscurs» il s'est peut-être produit bien des choses échappant à notre contrôle et à quoi seul le hasard donne l'aspect de l'archaïsme. Un exemple en forme d'avertissement est fortuitement fourni par le nom de nombre «dix» (ci-dessus 2.1.2.2) qui, dès l'indo-iranien, a dû être **dáća*, avec **-á-* < **-é-* (cf. latin *decem*). Si dans une des langues du Nouristan, le Prasun, apparaît *lεz'e* «dix», cet *-ε-* n'est sûrement pas une réplique de la forme phonétique indo-européenne, mais, comme l'a écrit le savant norvégien Georg Morgenstierne, le grand pionnier de la recherche en Nouristani, un «mirage scientifique»; cela vaut également pour *sëtë* «sept» en Prasun, dont la consonance avec lat. *septem* (mais

véd. *saptá* = avest. *hapta*, cf. ci-dessus 1.1.3.1) doit être classée parmi les « feux follets ».

2.2.

Quelles conclusions a-t-on tirées des apports des langues du Nouristan décrites ci-dessous ? De fort diverses. Un groupe de chercheurs voit dans ces langues, qui par leur évolution phonétique se rapprochent de l'iranien, et de l'indo-iranien par leur vocabulaire, le reste d'un troisième groupe de l'« aryen », à côté de l'indo-aryen et de l'iranien, comme il a été déjà mentionné (ci-dessus 2) ; cette conception serait particulièrement importante pour une présentation des dialectes indo-européens. Il y a pourtant d'autres modèles qui tentent d'interpréter les relations spéciales du Nouristani avec l'indo-aryen d'une part, l'iranien de l'autre. Ainsi quelques chercheurs ont admis que le Nouristani se serait séparé de la ligne évolutive indo-iranienne/indo-aryenne plus tard que l'indo-iranien mais plus tôt que le plus ancien védique ; d'autres voient dans la Nouristani une langue originairement identique à l'iranien, qui ensuite en fut séparée et passa sous l'influence secondaire de l'indo-aryen (ce qui semble s'accorder avec l'accord ancien intervenu avec l'iranien dans l'évolution phonétique ainsi qu'avec l'accord intervenu plus tard avec l'indo-aryen dans le vocabulaire). Tous ces modèles ont en commun qu'ils jugent impossible d'expliquer les langues du Nouristan simplement comme indo-aryennes (tandis que le groupe des langues dardiques des régions montagneuses nord-ouest du domaine linguistique indo-aryen, auquel appartient la langue archaïque de la littérature Kaśmīrī, doit être classé comme indo-aryen), ce qui, en tout cas, souligne la spécificité des langues du Nouristan dans le cadre de l'indo-iranien. Il a été indiqué ci-dessus (2.1.2.2.1) pourquoi il n'est pas possible de définir encore plus nettement la place des langues du Nouristan : leur attestation ne commence qu'au XIXe siècle de notre ère tandis que l'indo-aryen de Mittani est datable des XIVe-XVe siècles avant J.-C. (ci-dessous 3.2). C'est à bon droit que Georg Morgenstierne a déploré que le territoire du Nouristan ait probablement fait partie de l'empire Maurya (IIIe siècle av. J.-C.), sous l'empereur Aśoka, lequel fit composer, en diverses langues indigènes, ses inscriptions de propagande pour une loi morale générale (*dharma*) ; mais malheureusement, Aśoka n'a pas tenté de convertir les peuples du Nouristan par des inscriptions conçues en leur langue. Nous serions nettement plus avancés si avaient été conservées quelques lignes de Nouristani Aśokien !

3

Pour les questions que pose la linguistique indo-européenne, les plus anciennes langues indo-aryenne et iranienne dont la notation remonte au second millénaire avec J.-C. sont naturellement d'une importance capitale.

3.1.

Indo-aryen ancien. Cette notion coiffe le védique (3.1.1) et le sanskrit (3.1.2). Il est connu par une immense littérature. Comme dans le cas d'autres grandes langues littéraires « mortes », par exemple le grec ancien ou le latin, nous en savons sur l'indo-aryen ancien presque autant que pour une langue « vivante » susceptible d'être étudiée sous le rapport de ses multiples constituants : presque tout ce qui a été lexique, grammaire, syntaxe en indo-aryen ancien est consigné dans l'une ou l'autres des œuvres littéraires rédigées dans cette langue. Ici réside un contraste frappant avec la tradition de l'iranien ancien (cf. 3.3.3).

3.1.1.

La forme la plus ancienne de l'indo-aryen ancien est la langue de cette grande littérature qu'on appelle Veda au sens le plus large ; cet indo-aryen le plus ancien, subdivisé en niveaux d'ancienneté, est appelé *védique*. Le Veda consiste en quatre recueils de chants (*Saṁhitā's*) parmi lesquels la collection du Rig-Veda (*ṛg-veda, ṛk-saṁhitā*), par ses livres les plus anciens, est le monument littéraire indo-aryen le plus archaïque. Le Rig-Veda contient à lui-seul 1028 hymnes, en tout plus de 10 000 strophes ; à chacune des quatre *Saṁhitā's* se rattachent des *Brāhmaṇa's* (amples manuels de savoir sacrificiel), des *Āraṇyaka's* et des *Upaniṣad's* (textes traitant de mystique sacrificielle et de philosophie sacerdotale), enfin des textes concernant les sciences auxiliaires du Veda : les *Sūtra's*. Si l'on considère que la seule édition moderne des traductions de la plupart des divers Brāhmaṇa's occupe plusieurs volumes, on se fera une idée de la dimension des écrits notés en langue védique.

3.1.2.

Une forme de la langue haute indo-aryenne telle qu'elle avait été atteinte au milieu du 1[er] millénaire avant J.-C. – le védique ancien se distingue par ses archaïsmes grammaticaux, son système d'accents ancien, voire par une partie de son vocabulaire qui s'éteint après le Veda – est codifiée par la grammaire de *Pāṇini* (environ 400 av. J.-C.), une des plus grandes performances linguistiques de l'humanité ; la norme linguistique ainsi fixée deviendra, sous le nom de sanskrit (=*saṁ-skṛta-* « préparé, élaboré » [pour être

utilisé comme une langue sacrée]), l'organe d'une gigantesque littérature dans laquelle figurent les grandes épopées (le seul *Mahābhārata* contient en gros cent mille doubles vers !), la poésie classique tant dramatique que lyrique, plus une masse d'ouvrages scientifiques (philosophie, grammaire, lexicographie, médecine, science naturelles, mathématiques, astronomie, astrologie, etc.). Le sanskrit est resté vivant jusqu'à nos jours, tel le latin dans certains centres d'érudition occidentale, comme langue vivante ; en assistant à la Conférence internationale du sanskrit de Bénarès (1981), on put entendre des controverses passionnées de savants indiens s'exprimant dans un sanskrit parfait.

Des évolutions plus récentes de l'indo-aryen surgissent pourtant de bonne heure. L'empereur Aśoka (ci-dessus 2.2) fait établir ses édits en des langues épichoriques moyen-indo-aryennes ; les jeunes religions réformées du milieu du premier millénaire avant J.-C., le bouddhisme et le jaïnisme, se servent largement de langues moyen-indo-aryennes ; dans les drames de la littérature indienne classique, seuls les brahmanes et les rois parlent sanskrit, tandis que tous les autres personnages usent de diverses formes du moyen-indo-aryen (*Prākrit*). Vers le XII[e] siècle chrétien apparaît la forme linguistique néo-indo-aryenne ; de nos jours, une grande partie du sous-continent indien parle des langues néo-indo-aryennes, parmi lesquelles d'importantes langues littéraires comme le Hindī et le Bengālī. En dehors de l'Inde, un groupe de langue néo-indo-aryennes rencontre un vif intérêt : les langues des Tziganes asiatiques et européens.

3.1.3.

Pour abondante que soit notre information concernant la forme la plus ancienne de l'ancien-indo-aryen (3.1.1), nous n'avons pas de données exactes pour le début chronologique des textes ancien-indo-aryens ; il existe seulement des estimations. L'opinion répandue selon laquelle la rédaction du recueil d'hymnes du Rig-Verda se situerait entre 1200 et 1000 av. J.-C., tandis que tels hymnes anciens pourraient dater de plus tôt, n'est sans doute pas très éloignée de la vérité ; mais cela n'a rien d'une datation précise. Une circonstance d'autant mieux venue fut qu'au cours des cent dernières années un témoignage de la langue-indo-aryenne, localisé loin à l'Ouest de l'Inde, en Asie Mineure a fourni une lumière croissante et permis le rattachement aux chronologies exactes d'Égypte et de Mésopotamie. Le fait qu'on soit en présence d'une forme assurée du rameau indien de l'aryen (indo-iranien), mais attestée loin de l'Inde, constitue une des raisons de parler d'une langue « *indo-aryenne* » et non simplement « *indienne* » ; la raison la plus importante est sans doute qu'en Inde (Antérieure) sont encore localisées d'autres langues non indo-européennes importantes, surtout la famille des langues dravidiennes (avec le Tamil, vieille langue littéraire) et le groupe des langues Munda. Il est donc inadéquat et injuste, dans

ces circonstances, de réserver l'étiquette d'indiennes aux seules langues indo-européennes (cf. aussi ci-dessous 3.3.2.1).

3.2.

Cette attestation ancienne, exactement datable de la langue indo-aryenne en dehors de l'Inde à laquelle nous faisions allusion ci-dessus, est en rapport avec des restes linguistiques indo-aryens en relation avec l'éphémère royaume de *Mittani*; en peu de générations, approximativement entre 1400 et 1330 avant J.-C., cet empire a en tout cas joué, en Mésopotamie, un rôle de grande puissance en coexistant pacifiquement avec l'Égypte des temps glorieux d'Aménophis III et de ses prédécesseurs, avant de succomber à l'extension de l'empire hittite. Il importe de rectifier le mythe selon lequel l'empire de Mittani aurait été «une grande puissance aryenne», né d'une motivation politique qui a prévalu pendant une certaine période de l'histoire récente. C'était avant tout un État des *Hourrites*, peuple important dans l'histoire de l'Orient ancien, parlant une langue qui lui est propre, ni sémitique ni indo-européenne. La preuve de cet élément indo-aryen inhérent à la culture hourrite consiste seulement en un certain nombre d'anthroponymes et de théonymes et un groupe d'expressions techniques de l'entraînement hippique (3.2.1-3.2.3). Pour expliquer la provenance de ces résidus linguistiques indo-aryens dans un environnement hourrite, diverses hypothèses ont été proposées; on se rangera, de préférence, à l'opinion du spécialiste allemand du hourrite, Gernot Wilhelm, dont la compétence égale la prudence; on peut accorder créance au modèle «selon lequel des groupes parlant indo-aryen se seraient séparés du courant principal des tribus se dirigeant vers l'Inde en passant par l'Iran et finalement seraient entrés avec les Hourrites dans la zone d'attraction du Croissant fertile. Ils furent ainsi rapidement assimilés par le milieu hourritophone, marque par les cultures mésopotamienne et syrienne, si bien qu'il est possible qu'ils aient perdu très tôt leur langue d'origine». Les reliques linguistiques indo-aryennes de Mittani sont donc plus importantes pour le linguiste que pour l'historien; répondant à l'incertitude qui entoure une définition exacte des débuts de l'indo-aryen en Inde, nous sommes ici en présence d'une forme de l'ancien-indo-aryen précédant de peu le plus ancien védique, et qui peut être mis en rapport avec les règnes des pharaons Thutmoses IV (1400-1390 av. J.-C.), Amenophis III (1390-1352 av. J.-C.) et Echnaton (1352-1336 av. J.-C.) ainsi que du grand roi hittite Suppiluliuma I (environ 1355-1320 av. J.-C.), forme donc qui peut se situer aux XVe-XIVe avant J.-C. Nous ne possédons de données chronologiques aussi précises pour aucune autre langue indo-iranienne; les estimations acceptables pour les débuts de l'indo-iranien en Inde et en Iran ne remontent pas aussi loin dans le passé (ci-dessus 3.1.3, ci-dessous 3.3.1).

L'approche la plus stimulante des restes linguistiques de Mittani consiste en une esquisse de l'histoire de la science : comment, petit à petit, en relation avec les découvertes archéologiques célèbres et leur exploitation, ces reliques de l'indo-aryen le plus ancien sont-elles venues à notre connaissance ?

3.2.1.

Au début : embarras et erreurs. Les archives cunéiformes mises au jour en 1889 à *El Amarna*, site moyen-égyptien, contenaient la correspondance en langue surtout akkadienne d'Amenophis III et de son successeur Amenophis IV (Echnaton) avec les souverains d'Etats orientaux, parmi lesquels les rois de Mittani qui avaient même contracté des alliances dynastiques avec ces pharaons. On s'avisa dès le XIX[e] siècle que les noms de rois mittaniens tels que *Artatama* ou *Artaššumara* évoquaient des noms de l'Empire perse tels que *Artaxerxes*, *Artabazos* ; mais la seule interprétation qu'on put donner de l'existence de ces noms en Mésopotamie fut qu'on observait ici, en gros mille ans avant les rois de l'empire perse (par exemple Artaxerxes I, 465-424 av. J.-C.) la présence de Proto-Iraniens.

3.2.2.

La condition d'une analyse plus poussée fut une découverte ultérieure faite encore dans une métropole et dans des archives : le dégagement de la capitale de l'empire hittite et de ses archives officielles opéré après 1905 par l'assyriologue allemand Hugo Winckler ; le lieu de la trouvaille est le village turc de *Boğazköy* (depuis 1937 Boğaskale), à 150 km au nord-est d'Ankara. Cette découverte ne fut pas aussitôt exploitable en philologie, car ces archives contenaient surtout des textes en hittite, langue encore incompréhensible ; seul un petit nombre de documents était rédigé en akkadien (assyro-babylonien), langue internationale de la diplomatie, et que l'assyriologie avait déjà déchiffrée ; s'y trouvait un traité passé entre le grand roi hittite Suppiluliuma et un roi de la décadence mittanienne, Sattiwaza, qui devait, par ce traité, se plier aux conditions du roi hittite : ce traité, présenté dès 1907 à l'opinion scientifique par Hugo Winckler, contenait une donnée sensationnelle : dans une longue énumération de dieux du serment à la fin du traité (en rang médiocre, de la cent cinquième à la cent huitième place, donc en qualité de dieux « annexes » d'une maison royale qui, dans sa religiosité vivante, honorait de longue date des dieux hourrites) étaient cités quatre théonymes qu'on connaissait bien par le Rig-Veda : *Mitra* et son compagnon *Varuṇa; Indra* ; enfin les jumeaux divins (« Dioscures »), les deux *Nāsatyā*.

Si mince qu'ait pu être l'importance de ces dieux dans la vie contemporaine du Mittani, leur simple citation montrait qu'un élément indo-aryen était amalgamé au hourritisme mittanien. Il n'était plus possible, désormais, de songer sérieusement à des Iraniens archaïques (cf. 3.2.1) : parmi

ces dieux, seul *Mitra* (avest. *Miθra*) est demeuré vivant en Iran; les noms d'*Indra* et des *Nāsatyā* ont pour correspondants d'obscurs noms de démons dans l'Avesta (où le correspondant des *Nāsatyā* est une idole unique, non un couple de jumeaux); il manque en Iran un théonyme à rapprocher de *Varuṇa*. Il est intéressant pour l'histoire des sciences de rappeler ce que le génial et éclectique antiquaire Eduard Meyer – qui, peu de temps auparavant, avait considéré les noms princiers du domaine mittanien comme des « attestations les plus anciennement datées de la langue iranienne » – déduisait de ces théonymes (1909) : pour la première fois, la période indo-iranienne (aryenne) postulée par la recherche « se présente à nous par un document ». Aujourd'hui, l'opinion a prévalu que précisément ces théonymes spécifiques – de même que des éléments de vocabulaire identifiés plus tard (ci-dessous 3.2.3.) – permettaient de reconnaître la forme indo-aryenne de l'indo-iranien; que, je le répéte, les porteurs de cette langue étaient « des groupes indo-aryens » qui « s'étaient séparés (ci-dessus 3.2.) du courant principal des tribus en marche vers l'Inde »; ce qui importe au linguiste qui rédige ces pages, c'est que l'histoire de l'indo-aryen – loin de l'Inde – est ainsi appréhendée en un moment datable avec précision (ci-dessus 3.1.3.)

3.2.3.

Nous sommes également redevables aux archives de Boğazköy de la plus proche information concernant l'indo-aryen de Mittani. Dès la première campagne de fouilles (1906-1912), avait été exhumée une œuvre du XIVe siècle avant J.-C. dont l'auteur était un entraîneur de chevaux « du pays de Mittani » : *Kikkuli*. C'est un traité de l'entraînement et des soins à donner aux chevaux de charrerie, écrit en un hittite fautif, par un Mittanien, à l'intention du public hittite; l'exploitation de ce texte devint possible après que le déchiffrement du hittite fut acquis en 1915. Le fait que le spécialiste venu de Mittani ait écrit, pour les Hittites, ce manuel technique où apparaissent des *termini technici* suggère que les Mittaniens avaient une solide expérience de l'élevage et de l'entraînement des chevaux : ces connaissances étaient associées à l'emploi du char de combat à deux roues, dans une technique de guerre qui avait eu sans doute une part importante dans l'expansion de l'Etat mittanien. Certes, on est revenu aujourd'hui de l'opinion selon laquelle le char de combat à deux roues serait lui-même une invention des Indo-Aryens qui l'auraient importé en au Proche Orient; des recherches récentes ont montré clairement que le char de combat à deux roues est l'aboutissement d'une assez longue évolution intervenue en Asie Mineure. Les Indo-Aryens de Mittani ont sans doute combiné leur familiarité avec l'élevage et la maîtrise du cheval – qu'ils partageaient avec d'autres peuples indo-européens anciens et qu'au témoignage de leur onomastique et de la phraséologie de leur poétique, ils estimaient très haut – avec une technique du char de combat trouvée sur place, en Asie

Mineure, et ce avec assez de succès pour avoir été sollicités, au temps de la décadence mittanienne, comme experts hippologiques, par les Hittites plus heureux.

Les *termini technici* indo-aryens mentionnés sont les désignations de différentes distances de course ; ils associent un terme second °*wa-ar-ta-an-na* « tour de piste » (cf. védique *vart-* « tourner, virer », lat. *uertō*) aux noms de nombres impairs de l'indo-aryen : *a-i-ka*° « un » (védique *éka-*, plus ancien **áika-*), *ti-e-ra*° « trois » (védique *trí-*) ; *pa-an-za*° « cinq » (védique *páñca*), *ša-at-ta*° « sept » (védique *saptá-*) et **na-wa*° « neuf » (dans **na-wa-wa-ar-ta-an-ni* « dans neuf jour » avec l'haplologie °*wa-wa*° > °*wa*°- : *na-wa-ar-ta-an-ni* ; védique *náva-*). Cet ensemble de noms de nombre bien connus est particulièrement convaincant étant donné que « un » est l'une des rares divergences dialectales entre l'indo-aryen et l'iranien (indo-aryen **aika-* : iranien **aiva-* « un » < indo-européen **oi-ko-* à côté de **oi-u̯o-* [et **oi-no-* > ancien latin *oino-* = lat. *ūnus* « un »]) : *aika*° dans le traité de Kikkuli, de même que la forme particulière des théonymes dans le traité avec Sattiwaza (3.2.2.), plaide pour le caractère *indo*-aryen des restes linguistiques indo-iraniens de Mitanni. Quelques autres appellatifs identifiés comme indo-aryens dans d'autres documents de la sphère du Mittani confirment cette constatation : la langue indo-aryenne du Mittani en liaison avec la maison royale et l'aristocratie charrière du Mittani sans doute abandonnée rapidement en faveur du hourrite, était pour ainsi dire identique à une forme archaïque du védique.

3.3.

Nos connaissances de l'*iranien ancien* se limitent à deux « Corpora » de textes : l'*Avesta*, partie conservée d'une littérature en relation avec l'action du fondateur de religion *Zaraθuštra* (Zoroastre) (3.3.1), et les inscriptions royales en *vieux perse* de la dynastie des « *Achéménides* » (3.3.2.).

3.3.1.

La recherche n'est pas encore parvenue à s'accorder sur le lieu d'origine et les dates de vie de Zaraθuštra. L'on peut toutefois accréditer que son action eut pour théâtre l'Iran oriental ; et, parmi les datations de sa vie, la plus vraisemblable le place au XIe siècle avant J.-C. Étant donné que les parties les plus anciennes de l'Avesta – qui sont en même temps les plus anciens textes iraniens – sont ces recueils de chants religieux qu'on appelle *Gāθā's* et qui procèdent de Zaraθuštra lui-même, l'iranien serait attesté depuis 1 000 ans environ avant J.-C. ; cet âge correspond à celui que les estimations attribuent à l'indo-aryen en Inde (ci-dessus 3.1.3). Les reliques d'indo-aryen datables sans équivoque dans le royaume mittanien d'Asie Mineure sont plus anciennes (ci-dessus 3.2).

3.3.1.1.

La langue iranienne ancienne de l'Avesta est appelée *avestique* d'après la littérature qu'elle note ; on distingue deux dialectes : vieil-avestique (qui est limité aux *Gāθā's* et à un petit nombre d'autres textes) et avestique récent (dans lequel sont rédigés d'autres précieux textes anciens, mais qui, à l'inverse du vieil avestique, a été aussi utilisé pour la rédaction de textes récents parfois fautifs). On appelle simplement « avestique » la langue de l'Avesta pour la raison que cette langue est-iranienne ne peut être mise en relation certaine avec aucun des peuples iraniens connus ; c'est comme si la langue de l'Ancien testament ne pouvait être définie comme hébréo-araméenne et qu'on dût l'appeler « Ancien-Testamentaire ».

3.3.2

En revanche, il est possible de déterminer sans ambiguïté le lieu et le temps du *vieux perse*. C'est l'ancienne langue de la *Perse*, donc un dialecte sud-ouest-iranien ; il doit être mis en relation avec la dynastie achéménide, surtout avec les Rois des Rois Darius 1er (522-486 av. J.-C.) et Xerxès (486-465 av. J.-C.), son sucesseur. La tradition du vieux perse, dans une écriture cunéiforme créée spécialement pour cette langue, commence par la première inscription de Darius (vers 520 av. J.-C.).

3.3.2.1.

Les inscriptions des Achéménides sont pour une grande part trilingues : à côté de la version en iranien vieux-perse, elles présentent une version en langue élamite et en langue babylonienne. L'Iran ancien était un État pluri-ethnique dont l'élément indo-européen iranien ne formait qu'une partie. La vieille langue de culture qu'est l'élamite, attestée depuis le XXIIIe siècle et associée à un État important de l'Iran ancien, l'empire d'Elam, est un facteur essentiel de l'histoire linguistique et culturelle de la sphère iranienne ; jusqu'à nos jours, nous trouvons des langues non indo-européennes en Iran et de langues iraniennes hors d'Iran (par exemple dans l'ex. URSS et au Pakistan). Par analogie avec « indo-aryen » (ci-dessus 3.1.3.), on pourrait admettre un terme « irano-aryen », que recommande l'iraniste français Gilbert Lazard, et cela entre autres « parce qu'il respecte la sensibilité des iranophones non iraniens ». Cette désignation raisonnable ne s'est pas imposée pour l'instant.

3.3.3.

Le fait que l'iranien ancien ne soit attesté que dans les inscriptions en vieux perse et dans l'Avesta fait une différence importante avec l'ancien-indo-aryen ; tandis qu'ici nous sommes en présence d'une immense littérature, issue de nombreux siècles, qui renferme sensiblement la totalité du

stock linguistique indo-aryen ayant existé (ci-dessus 3.1.), les attestations de l'iranien ancien sont nettement limitées. Les inscriptions en vieux perse ne font qu'un mince corpus de textes; de plus, il s'y trouve maintes répétitions formulaires par lesquelles le stock linguistique effectif est encore appauvri. L'Avesta est certes plus ample, mais ici encore existent des restrictions thématiques, à savoir que de nombreux contenus ne sont jamais exprimés dans ces textes, si bien que diverses formes que possédait la langue n'y sont pas transmises. Le matériel attesté ne représente qu'une fraction de ce que furent le vieux perse et l'avestique quand ils étaient encore vivants et parlés. Mais plus grave encore est la circonstance suivante : nous n'avons de documents que de l'avestique est-iranien et du vieux perse sud-ouest-iranien pour l'époque iranienne ancienne. Les nombreux dialectes iraniens qui nous sont connus dans les périodes moyen-iranienne et néo-iranienne, par exemple les langues des Sogdiens, des Parthes, des Chorasmiens, des Saces, etc., ont sûrement existé dès l'époque iranienne ancienne; ils ne nous sont pas parvenus. Il en résulte un problème particulier à la linguistique indo-iranienne : faire apparaître les éléments de ces langues qui ont existé, mais ne sont pas documentés, dans les « traditions adventices » (3.3.3.1.), ainsi que dans les langues iraniennes récentes qu'atteste l'héritage indo-iranien et indo-européen ancien existant naturellement à l'époque lacunaire de la tradition de l'iranien ancien, mais non attesté dans les textes originaux de cette langue (3.3.3.2.).

3.3.3.1.

Grâce à l'importance politique de l'antique Empire des Perses, des mots et des noms tirés de langues iraniennes anciennes nous ont été transmis dans divers textes non iraniens, et, parmi eux, nombre de mots et de noms qui n'apparaissent pas dans les textes iraniens anciens eux-mêmes; on parle ici d'une « tradition indirecte » de l'iranien ancien. La plus connue parmi les auteurs grecs est la tradition adventice d'Hérodote qui transmet aussi des mots et noms religieux d'une langue iranienne ancienne, le *scythe*, non attestée par ailleurs, ainsi qu'une série de termes vieux perses et mèdes qui manquent dans les sources originales de l'iranien ancien; le plus frappant, c'est qu'un terme technique aussi connu de l'Empire perse que « satrape » (grec σατράπης) n'apparaît que sous la forme qui fut courante à l'ouest, donc d'une « tradition adventice » (on reconstruit un mède **xšaθra-pā-* « protégeant l'empire »). L'importance de la tradition adventice grecque est surpassée par la tradition adventice en langue élamite (ci-dessus 3.3.2.1.); ainsi les termes élamites de la comptabilité de Persépolis renferment des termes techniques iraniens et surtout quelque 2 000 noms propres issus des différents dialectes iraniens anciens de l'État achéménide que la tradition iranienne originale ignore. Nous avons surtout une bonne idée de la langue des *Mèdes* bien que jusqu'à

présent aucun texte médique n'ait été mis au jour; cette connaissance du mède provient aussi du fait que manifestement, dans les inscriptions royales en vieux perse, certaines expressions sont empruntées au modèle de la langue mède. En voici un seul exemple : quand Darius Ier dit des provinces de son empire qu'elles « contiennent toutes les tribus », il emploie à cet effet un mot composé tiré des correspondantes iraniens du védique *víśva-* « tout entier » et *jána-* « peuple, tribu ». Selon les lois phonétiques du vieux perse, on attendrait **visa-dana-*; or l'inscription en vieux perse de Darius Ier a *vispa-zana-* ce qui représente une évolution « noble » mède avec *-sp* = véd. *-śv-*, *z* = véd. *j*. Dans la version élamite de son inscription trilingue (3.3.2.1.), on trouve cependant à cette place la transcription du mot **visa-dana-* dont on déduit le phonétisme authentique en vieux perse ! Ce qui revient à dire que les interprètes élamites avaient entendu **visa-dana-* de la bouche des Perses, tandis qu'eux-mêmes, dans la version qui rend la langue maternelle du roi des Perses, emploient l'expression médique, plus chargée de tradition.

3.3.3.2.

L'attestation restreinte de l'iranien ancien – deux dialectes seulement, et une quantité limitée de textes – rend nécessaire le recours aux langues moyens-iraniennes et néo-iraniennes pour l'explication de l'indo-iranien et de l'indo-européen; la situation diffère donc de celle qui prévaut dans les domaines abondamment documentés de langues anciennes telles que l'indo-aryen ancien, le grec ancien, le latin, où il est rare qu'on doive invoquer les niveaux moyens et récents pour la reconstruction. En voici quelques exemples :

(1) Au védique *tántra-* « trame d'un tissu » correspondrait l'iranien ancien **tanθra-* : mais ni l'Avesta ni les inscriptions achéménides ne fournissent ce terme pour « trame d'un tissu ». Cependant **tanθra-* est phonétiquement l'archétype de néo-perse *tār* « trame d'un tissu, fil », afghan *tōr* « filet, lacet », waxi (une des langues du Pamir) *tōr* « filet », l'indo-iranien **tantra-* (apparenté à lat. *tentus* « tendu ») n'est donc expliquable du côté iranien qu'à partir de langues d'aujourd'hui.

(2) Védique *lopāśá-* « renard, chacal » aurait pour appui en iranien ancien **raupāsa-*; le mot n'apparaît dans aucun texte iranien ancien. Mais il est largement appuyé dans les langues moyen- et néo-iraniennes. Ainsi en parthe *rwb's*, ossète *robas* « renard » et d'autres. Dans les listes de comptabilité élamites de Persepolis apparaît toutefois un nom iranien ancien *Ra-u-ba-sa*; cet homme s'appelait manifestement « Renard », néo-perse *Rūbah*, « Renard » comme nom propre. Ici se montre notre seconde source pour les éléments linguistiques iraniens anciens absents des sources originales : la tradition indirecte (ci-dessus 3.3.3.1.).

(3) Védique *snuṣā́-* « belle-fille », est un vieux mot indo-européen (russe *snoxá*, lat. *nurus*, dans le haut-allemand de Martin Luther encore *Schnur* « belle-fille », etc.) et n'est pas attesté dans les langues iraniennes anciennes, mais apparaît en sodgien et dans plusieurs langues néo-iraniennes. Ainsi la langue waxi du Pamir citée ci-dessus a un mot *sitəx̌* « belle-fille » que l'art du comparatiste sait ramener à l'ancien-iranien **snušā-*.

Ces exemples pourraient être aisément multipliés. Mentionnons encore seulement qu'il y a des cas où des termes indo-européens n'existent ni en indo-aryen ni en iranien ancien mais apparaissent dans une langue néo-iranienne. Ainsi le terme important pour l'histoire culturelle désignant le « saumon » (russe *losós'*, all. *Lachs* « saumon », tokh. B *laks* « poisson », etc.) n'apparaît pour désigner un poisson qu'en ossète dans le domaine indo-iranien : *läsäg* « saumon ».

3.4.

Ce dernier exemple a déjà suggéré que les langues iraniennes peuvent fournir des indications de nature indo-européenne auxquelles l'indo-aryen n'a pas part. L'opinion de maint ancien indo-européaniste selon laquelle les témoignages de l'indo-aryen ancien suffisait à la science indo-européaniste – grâce à ses vieilles, surabondantes attestations textuelles transmises fidèlement par des génies phonéticiens comme les Indiens anciens –, cette opinion est aujourd'hui fondamentalement dépassée ; les langues iraniennes, selon cette opinion, avec leur tradition lacunaire, phonétiquement fruste, ne seraient qu'un démarquage de l'indo-aryen ancien et n'intéresseraient que les iranistes, non les indo-européanistes. Ce point de vue méconnaît que l'iranien a conservé des archaïsmes qui manquent en indo-aryen. Je citerai seulement trois exemples à l'appui de l'affirmation selon laquelle, dans le composé « indo-iranien », les *deux* termes importent à une perspective indo-européenne (ce qui marquera la conclusion rationnelle de cette présentation de l'indo-iranien).

(1) Dans le nom du « père », la syllabe initiale avait une laryngale, consonne qui, dans certaines conditions (développement d'une voyelle épenthétique) se vocalisa en *i*, dans d'autres conditions s'amuït en zéro (Ø) : indo-européen *ph_2tér-m̥* « le père » (accusatif) ; *ph_2tr-éi* « au père » (datif) (de *ph_2tér-* = lat. *pater*, etc.). On attendrait en indo-iranien **pitár-am*, **p(Ø)tr-ái*. En védique, la forme *pit(a)r-* a été, par analogie, généralisée à tous les cas : *pitáram*, *pitré*. L'avestique en revanche continue l'ancienne alternance *i/Ø* : *pitarəm* « le père » (acc.) mais **ptrai* > **fθrai*, noté *fəδrōi* « au père » (dat.)

(2) Les radicaux du type lat. *gen-us*, *gen-er-* avaient une apophonie du suffixe : *-os-* (proto-latin **gen-os* > *genus*) opposé à *-es-* (proto-latin **gen-es-a* > *genera*). En indo-aryen l'indo-européen *-e-* (avant de devenir *-a-*, cf. ci-dessus 1.1.4.1) provoque la palatalisation d'un *-g- précédent en

-*ǰ*- (ci-dessus 2.1.2) ; devant indo-européen *o*, -*g*- reste -*g*-. Un proto-indo-iranien **aug-os*/*aug-es*- « force » (cf. lat. *augustus* « majestueux ») devrait donc fournir en indo-iranien la variation **augas*/**auǰ-as*- ; mais en védique c'est la seule variante **auǰas*- (ved. *ójas*-) qui apparaît généralisée par analogie : *ójas* « la force », *ójas-ā* « par la force ». L'ancien-avestique a conservé l'alternance -*g*-/-*ǰ*- qui reflète le jeu de l'apophonie -*os*/-*es*- : « la force » *aogō*, *aoǰaŋhā* « par force ».

(3) Enfin, l'information suivante est d'une importance décisive pour l'indo-européanistique moderne : le mot désignant le « chemin, sentier », est en proto-indo-iranien un dérivé à suffixe à laryngale -*oh*$_2$- de la racine **pent*- (cf. russe *put'* « chemin » [∼*s-put-nik* « compagnon de route, étoile qui accompagne, satellite artificiel ; spoutnick » !] de **pont*-) ; nominatif singulier **pént-oh*$_2$-*s* « le chemin ». Devant un cas à accent sur la finale, les deux syllabes précédant la terminaison sont affaiblies : génitif singulier **pn̥t-h*$_2$-*és* « du chemin ». Étant donné qu'en indo-iranien la rencontre immédiate de -*t*- et -*h*$_2$- donne *th*, les deux formes devraient être **pántās*/**pathás*, avec une alternance de deux phonèmes différents (*t*/*th*) dans la syllabe radicale (*pant*-/*path*-). On ne saurait s'étonner si ces cas révèlent dès le plus ancien védique une assimilation en *th* dans la syllabe radicale : le védique a *pánthās* « le chemin », *pathás* « du chemin ». Cependant le produit de l'évolution phonétique attendue **pántās*/**pathás* est encore conservé en avestique : av. *pan̨tå*/*paθō* ne peut remonter qu'à cette forme. Dans ce cas, l'indo-iranien est le témoin-clé d'une approche complexe de l'indo-européanistique moderne, de la morphologie, de l'apophonie et la phonétique linguistique du type le plus archaïque de l'indo-européen.

Les deux dialectes si proches de l'indo-iranien contribuent donc ensemble, chacun à sa manière, au grand système des études indo-européennes. Émile Benveniste, le grand indo-européaniste français a, dans un autre contexte, trouvé une belle et valable formule : « le témoignage védique vaut par sa richesse, le témoignage avestique par sa fidélité ».

CHAPITRE 6

ANATOLIAN

Craig MELCHERT

1. INTRODUCTION

The Anatolian branch of Indo-European is attested in written documents from what is now Turkey and Northern Syria dating from approximately the sixteenth century B.C. to the first centuries A.D. To this group belong at least Hittite, Palaic, Luvian, Lycian and Lydian. I will argue below that we should add Carian, Pisidian and Sidetic, but the linguistic affiliation of these poorly attested and understood languages remains controversial, and in any case they can contribute little to our overall understanding of the Anatolian family. Unless stated otherwise, references to pan-Anatolian features in what follows are not meant to include them. It is unlikely that any of the Anatolian languages remained as a spoken language beyond the first centuries A.D., and there are no living representatives of the family.

The relationship of the Anatolian languages to the rest of Indo-European has been a matter of serious debate ever since the identification of Hittite as an Indo-European language by Bedřich Hrozný in 1915. The problem is complicated by the fact that the available evidence for Hittite

has far exceeded that for other members of the group. Hittite has hence been far more thoroughly analyzed and better understood. This imbalance has resulted in the peculiar situation that the debate over the relationship of Anatolian to Indo-European has been framed almost entirely in terms of Hittite, despite the fact that the close relationship of Hittite with the other members of the subfamily has been widely acknowledged for more than a half-century.

Dramatic new discoveries and further analysis have improved our understanding of the « minor » Anatolian languages to the point that some clarification regarding the status of the entire group is now possible. As I will show below, all members share certain highly characterized innovations which can only be explained by assuming a period of prehistoric common development. I call this stage for convenience « Common Anatolian », but I stress that this notion is primarily linguistic (on a par with Common Celtic or Italic), not geographic. I must note explicitly that Phrygian does not belong to this group, although it is an Indo-European language attested in Anatolia (see the separate discussion by Professor Brixhe in this volume). The existence of a common prehistoric source for the Anatolian languages as thus defined is independent of any hypotheses about population movements. The speech community associated with Common Anatolian may or may not have lived on the soil of Anatolia.

The evidence for a common prehistoric development falsifies the idea that the resemblances among some of the Anatolian languages are due merely to convergence or contact. However, it does not settle the issue of whether the rest of the Indo-European languages underwent a period of common prehistoric development apart from the Anatolian group. It is immaterial whether one frames this problem in terms of « Indo-Hittite », « early separation » of the Anatolian group, or « early » *vs.* « late Indo-European » features. The question remains: how can we best explain both the resemblances and differences between the Anatolian languages and the rest of Indo-European ? I can only allude to this issue here. In what follows, I will present Anatolian as a subgroup of Indo-European, on a par with Germanic, Slavic, etc. However, this should not be taken to prejudge the question just posed, which in my view remains open.

2. The Languages

2.1 Hittite

Hittite was the chief administrative language of the Hittite Empire. It is attested in extensive cuneiform documents dating from the sixteenth to the end of the thirteenth century B.C. Advances in scholarship over the past quarter-century now enable us to assign a relative chronology for most manuscripts and compositions, and we may distinguish Old Hittite (1570-1450 B.C.), Middle Hittite (1450-1380 B.C.), and Neo-Hittite (1380-1220 B.C.). This division of the corpus continues to undergo refinement, but the ability to differentiate older from newer manuscripts has already revolutionized study of the language. Nearly every aspect of Hittite grammar must be reevaluated in the light of this new knowledge.

The texts are inscribed on baked clay tablets (except for one splendid bronze example). The vast majority of tablets has been found at the Hittite capital of Hattusha (modern Boğazköy, now Boğazkale), in the middle of the central Anatolian plateau, but there have also been important finds at Maşat Hüyük to the northeast and at Ugarit (Ras Shamra) on the Syrian coast, with scattered discoveries elsewhere. A major trove of Hittite tablets just discovered at Ortaköy (south of Çorum) in north central Anatolia has yet to be fully evaluated.

The extant Hittite texts are quite varied in content, including historical texts such as annals and treaties, a « law code », and a series of instructions on the training of chariot horses. However, the bulk of the material concerns the state religion: cultic administration and ritual and related mythology.

Claims that Hittite has undergone some kind of « creolization » are unfounded, and the effects of « substrate » or « adstrate » languages on its development have also been grossly exaggerated. Hittite is unmistakably an Indo-European language in its grammatical structure and core vocabulary.

2.2. Palaic

Palaic was once the spoken language of the land of Pala, which in all probability lay to the northwest of the Hittite heartland, across the Halys River in classical Paphlagonia. It was certainly extinct by the thirteenth century B.C. and may have already been so by the time of our first texts in the sixteenth. We have only a few fragments of Palaic, preserved for cultic

purposes by the Hittites in their cuneiform archives in Hattusha. The very scanty evidence limits our knowledge of the language, but it is clear that it is an Indo-European language closely related to Hittite and Luvian, in some respects more conservative than either.

2.3. Luvian

We have evidence for two forms of Luvian. The Hittite archives at Hattusha contain a number of texts in «Cuneiform Luvian», with manuscripts dating from the Old to Neo-Hittite periods. While the number of manuscripts is reasonably large, recent research has shown that, aside from a few fragments, the extant material represents variations on only a handful of distinct compositions. As a result, the Cuneiform Luvian vocabulary known to us is severely limited, and evidence is lacking for some important grammatical categories. A number of Luvian words are also found scattered throughout the Hittite texts, both as genuine loanwords adapted to Hittite grammar and as unassimilated Luvian forms.

Given the close genetic relationship between the two languages, it is sometimes difficult to determine whether a given example in a Hittite context is an inherited Hittite word related to one in Luvian, or a borrowing from Luvian.

«Hieroglyphic Luvian» is attested in documents written in the so-called «Hittite hieroglyphs», a writing system apparently developed in Anatolia. The first extended texts in this system, monumental stone inscriptions dating from the thirteenth century, are very hard to interpret, because of the use of non-standard sign shapes and frequent «logographic» spellings which do not indicate the pronunciation. Most of our Hieroglyphic Luvian texts date from the tenth to seventh centuries B.C., after the fall of the Hittite Empire, and are thus much later than those for Cuneiform Luvian.

Aside from a few letters and economic documents written on soft lead strips (and personal seals with only names and titles), the Hieroglyphic Luvian texts are inscriptions on stone. Most are dedicatory in some sense, but many contain lengthy historical sections, and the Hieroglyphic Luvian corpus is more varied in content that that of any Anatolian language except Hittite.

The stone inscriptions are concentrated in southern Anatolia (classical Cilicia and Commagene) and northern Syria, but they are also found in central Anatolia (including Hattusha!) and as far west as modern Izmir. Luvian was certainly a major spoken language in large areas of southern and southwestern Anatolia during the Hittite Empire and afterwards. It

is likely that some form of Luvian was also once spoken across much of western Anatolia, and it is even possible, although unproven, that the language of the Trojans was Luvian (Watkins, 1986). To what extent and at what times Luvian was spoken in Hattusha and surrounding areas remains undetermined.

Major revisions in the reading of the hieroglyphic signs by Hawkins-Morpurgo Davies-Neumann (1974) have shown that Hieroglyphic and Cuneiform Luvian are very closely related forms of the same language. The older and highly restricted attestation of the latter makes it difficult to be more precise about their relationship. I follow others in treating them provisionally as parallel and independent dialects.

2.4. Lycian

We have somewhat less than 200 inscriptions on stone plus a number on coins in the native language of the kingdom of Lycia in southwestern Anatolia. These are written in an alphabet closely related to that of Greek. While a few coins may be older, the texts on stone all date from the fifth and fourth centuries B.C. Nearly all the texts are tomb inscriptions with fairly stereotyped contents. One notable exception is the «Inscribed Pillar of Xanthos», a dedicatory text with long historical portions, unfortunately still largely unintelligible due to problems of vocabulary. The «Létôon Trilingual», another dedicatory text in Lycian-Greek-Aramaic discovered in 1974, has dramatically improved our knowledge of Lycian.

Two extant texts from Lycia are written in a different dialect called Milyan or «Lycian B» (*vs.* «Lycian A», *i.e.* ordinary Lycian). While Milyan seems more conservative in a few respects, there is no justification for viewing it as an archaic form of ordinary Lycian. We should rather treat Lycian and Milyan as parallel independent dialects like the two forms of Luvian.

Lycian shares a number of innovative features with Luvian, leading some scholars to speak prematurely of a «Luvo-Lycian» subbranch of Anatolian. However, some of the relevant features are also shared with Lydian, while Luvian agrees with Hittite and Palaic in other respects. It thus seems more prudent at this point to view the Anatolian languages as a continuum of dialects.

2.5. Lydian

The language of the western Anatolian kingdom of Lydia is attested in over a hundred inscriptions, but only a few dozen of these are of significant length and reasonably complete. While some coins may date from as early as the eighth century B.C., the texts on stone are from the fifth and fourth centuries and thus roughly contemporary with our Lycian material. Likewise, the Lydian texts are written in an alphabet similar to that of Greek.

Most of the texts are dedicatory in some sense, and indeed mainly funerary. A few seem to be decrees of some sort. Remarkably, some are in verse, with a fairly regulated meter based on word accent and vowel assonance in the final words of each line (Eichner, 1986, with refs.).

A short Lydian-Aramaic bilingual permitted the first entry into the language, but we still lack a Lydian-Greek bilingual of any length. As a result, while we can parse most sentences with some assurance (except for some in the verse texts), our grasp of the meaning remains vague. This lack of precision in terms of the lexicon makes etymologizing particularly parlous, and we do not understand Lydian as well as any of the other languages cited above.

This is particularly unfortunate, because Lydian superficially appears much different from the other Anatolian languages. We cannot yet determine whether this is due to its having preserved archaisms lost elsewhere or to a series of unique prehistoric developments. In any case, Lydian clearly did share in certain crucial Common Anatolian innovations. Thus the notion that the appearance of Lydian in western Asia Minor might be due to a separate development from that of the other Anatolian languages is mistaken.

2.6. Carian, Pisidian and Sidetic

There are three more languages of southwestern Anatolia which we should probably add to the Anatolian group of Indo-European.

The language of Caria in the extreme southwest is attested in a few texts from Caria itself, probably dating from the fourth and third centuries B.C., but most evidence for the language comes from tomb inscriptions and graffiti of Carian immigrants in Egypt (seventh to fourth centuries B.C.). Decipherment of the Carian script has proven difficult and controversial. Ray (1981) has used Egyptian renderings of Carian names to establish

an objective basis for reading some signs, but his work has not yet won universal acceptance, and much work remains to be done in any case.

Evidence for Pisidian is extremely sparse: thirty-odd tomb inscriptions in alphabetic script from Northern Pisidia containing only names and patronymics. Precise dating is difficult, but second to third century A.D. seems most likely.

Sidetic, the language of Side, a city on the coast of Pamphylia, thus far appears in only a half-dozen inscriptions, all but one clearly dedicatory, dating from the third century B.C., and a few older monograms on coins, again written in an alphabet related to Greek.

My assignment of these problematic languages to the Anatolian family of Indo-European is based on a single highly characteristic construction: the same use of a «relational» adjective to mark possession which is also attested in Luvian, Lycian and Lydian (*i.e.* «paternal house» for «father's/fathers' house»). The Anatolian languages employ two suffixes in this function. The more common of these has the shape **-asso/ī-* (< PIE **-eh*$_2$*so/ih*$_2$*-*, as per Bader, 1985: 85, note 99, and others). For example, in Hieroglyphic Luvian /Halparuntiyas Laramassis nimuwitsas/ «Halparuntiya, son of Larama», the suffix appears in the animate nominative singular, modifying /nimuwitsas/ «son», which is in apposition to /Halparuntiyas/.

In Lycian original **s* becomes *h* between vowels and is lost at the end of a word. The above pattern thus is realized as: *Xudali Zuhriyah tideimi* «Xudali, son of Zuhri» (*-ah* < **-assos*). We similarly find in Pisidian *Musita Tas* «Musita, (son) of Ta». As in Lycian, both noun and adjective have lost the original final *-s* of the nominative singular, and the final unaccented vowel of the adjective has also been deleted, but the new final *-s* of *Tas* faithfully continues the original double *-ss-* of the adjectival suffix, like the *-h* in Lycian.

The same pattern recurs in Sidetic (*poloniw pordors* «Apollonius, (son) of Apollodoros») and Carian (*uksmu dkorś* «Uksmu, (son) of Dkor»), where the symbol *ś* indicates that the result of **-ss-* is somehow different in Carian from ordinary single *s*.

Lydian uses a suffix in *-l-* for the relational adjective: *eś wãnaś manelis alulis* «This tomb (is) of Mane, (son) of Alu» (confusingly, *ś* indicates an ordinary dental sibilant in Lydian, and *s* a palatal caused here by the preceding *i*). Carian also has examples of this suffix beside the *-ś* from **-asso/ī-*.

3. LINGUISTIC FEATURES OF THE ANATOLIAN LANGUAGES

3.1. Phonology

The three contrasting stop series of PIE are reduced to two in Common Anatolian (CA) by the merger of the « voiced aspirates » with the voiced stops: Hitt. *anda* « into » < **endo*, *andurz(a)* « inside » < **en-dhwr̥* + *« indoors ». The « functional load » of the remaining voiced/voiceless contrast is further reduced in CA by a series of conditioned changes. Voiced stops are generalized word-finally (cf. Old Latin): Hitt. pret. 3rd sg. *pait=as* [paydas] « went he » < **-t*, Lyd. *qid* « what », etc. Voiceless stops become voiced after accented long vowels (Eichner, 1973: 79f) and between unaccented vowels (Morpurgo Davies, 1982/3: 245-270, Eichner, 1986: 12-13, & Melchert, 1992a): Lyc. *tadi* « puts » < CA **dǽ́di* < PIE **dhéh$_1$ti* (cf. Grk. *títhēmi*); Lyc. *tubidi* « strikes » < CA **towbéyedi* < (virtual) PIE **(s)towbhéyeti* (cf. Grk. *stuphelízō*).

By the historical period all the attested languages have also generalized voiceless stops in word-initial position: Lyc. *tadi* « puts » (see above), Hitt. reduplicated *ti-tti-* « install » (also made from **dheh$_1$-* after the devoicing; on the geminate *-tt-* see below). In Lycian and Lydian voiceless stops merge with voiced stops after nasals: Lyc. *ñte* « into » < **endo* and pres. 3rd pl. *-Ṽti* < **-Vnti*, both pronounced [nd], likewise Lyd. *ẽt-* « into ». In Lycian voiced stops become fricatives in all other environments. Lydian also shows a strong tendency towards spirantization under complex and not yet fully understood conditions. On the other hand, in Hittite, Palaic and Luvian voiceless stops appear as geminates in intervocalic position (Hitt. *āppa* « back (wards) » < **ópV*; cf. Grk. *opíssō*), and it is likely that the voiceless/voiced opposition has been reanalyzed in the few environments where it remains as one of fortis/lenis.

The Anatolian languages are neither « centum » nor « satem », since Luvian and Lycian show that CA preserved the PIE three-way contrast of palatal, velar and labiovelar stops (Melchert, 1987 and 1989): PIE **k̂* > Luv. *z*/Lyc. *s* (CLuv. *zi-* = Lyc. *si-* « lie » < **k̂ei-*); **k* > Luv. *k* /Lyc. *k* (CLuv. *kīšāi-* « comb » < **kes-*, Lyc. *tuk(edri)-* « statue » < **twek-* *« body »); **k^w* > Luv. *kw*/Lyc. *t* before front vowel (Luv. *kwi-*/Lyc. *ti-* < **k^{w}i-* « who/what »). Palaic and Hittite apparently do merge the palatals with the velars. The Lydian evidence is too sparse for a determination.

Of the famous PIE laryngeals, **h$_1$* is nowhere directly preserved in Anatolian, despite some claims to the contrary. In CA **h$_2$* undergoes weakening under roughly the same conditions as the voiceless stops and appears in Hittite, Palaic and Luvian in a stronger from *-hh-* and a weaker

-h-, matched in Lycian by *x* and *g*. The sound is totally lost in Lydian. PIE $*h_3$ is surely lost everywhere except in word-initial position, where I share the view that it is preserved as *h-* in Hittite, Palaic and Luvian – others assume that is lost here too.

The PIE sibilant **s* is generally stable in Anatolian but for Lycian (A), where it is weakened to *h* except as the first member of a cluster. Cases of geminate *-ss-* resulting from assimilated clusters are common.

The PIE sonorants **m*, **n*, **l*, and **r* are for the most part maintained, again with geminates formed by various assimilations. The secondary vowel which develops from their syllabic forms is usually *a*, but in a few cases probably *u*. The glide **w* is almost universally preserved, but loss of **y* is widespread, especially between vowels.

The CA vowel system is virtually that of PIE, with the exception that the diphthong **ey* may be monophthongized to a close long **ẹ̄*.

However, the attested languages radically modify the CA system in complex ways which can only be alluded to here. All the languages merge **o* with *a* except Lycian, where **o* falls together with *e* (Melchert, 1992b). This leaves Hittite, Palaic and Lycian with basically four-vowel systems (*a, e, i, u*) – Luvian further reduces this to three by merging **e* with *i*. Lydian recreates a five-vowel system by adding a secondary *o* whose origins are generally unclear.

Lycian definitely and Lydian probably abandons any contrast between short and long vowels, but Hittite, Palaic and Luvian show massive lengthening of accented short vowels, a change whose details remain the subject of debate (Kimball 1983, Eichner, 1986: 13, Melchert, 1992c).

3.2. Nominal Morphology

The inflection of nouns and adjectives in Anatolian holds relatively few surprises from the Indo-European point of view. There are only two numbers, singular and plural – alleged traces of a dual in Hittite and Luvian are quite dubious. Old Hittite does distinguish productively an individual and collective plural: *e.g. lahhurnuzziyes* « branches » *vs. lahhurnuzzi* « foliage » to *lahhurnuzzi-* « leafy branch ». Luvian and Lycian also show traces of this distinction, comparable to Latin *locī* and *loca* to *locus* « place ».

It is now certain that Common Anatolian inherited three genders from PIE: masculine, feminine and neuter. The strongest evidence for the feminine is provided by Lycian animate nouns with *a*-vocalism such as *lada-* « wife » (Melchert, 1992b) and by the phenomenon of « *i*-motion » in at least Luvian, Lycian and Lydian (Oettinger, 1987, after Starke). However,

in the attested languages the feminine and masculine have merged into an animate class opposed to the inanimate (neuter).

There is evidence for nine cases in CA: vocative, nominative, accusative, genitive, dative, locative, allative, ablative-instrumental and ergative. The first six clearly continue PIE categories both functionally and formally. As expected, the vocative differs from the nominative only in the singular, and the neuter gender has a single form for the nominative-accusative in both singular and plural. The dative-locative has a single form in the plural, and the distinction of these two cases in the singular in CA is far from assured.

The vocative singular usually consists of the bare nominal stem. The Old Hittite vocative ending *-i* may or may not continue PIE **-e* of the *o*-stems displaced to other stem classes. PIE masculine/feminine nom. sg. **-s* is attested throughout Anatolian. There is a strong tendency to generalize the *-s* to paradigms where it was lacking in PIE (Hitt. *hāras* « eagle » *vs.* Lat. *homō* « man », Hitt. *hāssas* « hearth » *vs.* Lat. *arā*), but the extent to which this had been accomplished in CA cannot be determined. Masculine/feminine acc. sg. **-m* is reflected in nasal or nasalized endings across Anatolian. Likewise the nt. nom.-acc. sg. **-om* of *o*-stems versus the expected zero ending in this category in other stem classes. Genitive singular **-os* appears in Hittite, Palaic and Luvian *-as*, and Hittite has a few examples of *-s* < **-s* (*nekuz* [nek^{w}ts] « of evening »). Noteworthy is the archaic gen. sg. *-as* < **-os* in *o-* stems, attested in the same three languages. Hitt. dat.-loc. sg. *-i* probably continues PIE locative **-i*, while CLuvian dat.-loc. sg. *-ī* with a long vowel even when unaccented shows rather PIE dative **-ey*, but this is by no means certain. The matching *-i* ending of Palaic and Lycian could reflect either **-i* or **-ey*.

PIE masculine/feminine nom. pl. **-es* is maintained in Palaic *-es*, which also shows *-as* from the PIE *o*-stem ending **-ōs*. Hittite has *-ēs* with a long vowel in all classes, probably generalized from a contracted form of the *i*-stem ending **-eyes*. The Western languages Luvian, Lycian and Lydian have built a new animate nominative plural ending by adding an *-i* of uncertain origin to the old animate accusative plural. The latter, from PIE **-ns* (or better **-ms*) is most clearly preserved in Luvian *-nz(a)* and Milyan *-Ṽz* ([-nts] with epenthetic *-t-*). The PIE neuter (collective) nom.-acc. plural **-eh*$_2$ of the *o*-stems appears across Anatolian as *-a*, generalized to all stem classes. Old Hittite *āssū* « goods » is a precious archaism showing long final *-ū* < **-uh*$_2$ with original ending **-h*$_2$. PIE gen. pl. **-ŏ̄m* is preserved in Old Hitt. *-an*, Lyc. *-ẽ*, and Lydian *-aν* (in the last instance shifted functionally to a dative). Hittite and Palaic *-as* and Lycian *-e* continue a CA dative-locative plural in **-os*, comparable to Lat. *-bus*, OLith. *-mus*, etc. without the labial element. Luvian has a new ending *-anz(a)*, again after the animate accusative plural.

The remaining three cases cited above reflect CA innovations built on inherited PIE material. Old Hittite has an allative singular in *-a* indicating place to which (*aruna* «to the sea» *vs.* locative *aruni* «in/to the sea»). This ending, also attested in Palaic loc. sg. *-a*, continues PIE **-ō* (cf. Lat. *quō* «whither»). A CA ending **-(o)ti* (secondarily **-odi* in unaccented position) appears in Luvian *-ati*, Lycian *-e/adi*, and Hittite *-az/-z(a)* ([-ats/-ts]), the last with assibiliation of **t* before **i* and then apocope of the vowel. Indifferent to number, this case functions as both ablative and instrumental in Luvian and Lycian. In Old Hittite it has been restricted to ablatival function by the appearance of a new instrumental ending *-(i)t*. In later Hittite *-(a)z* ousts *-it* and regains both functions (for this sequence of events see Garrett, 1990: 278). CA abl.-inst. **-(o)ti* has comparanda in Tocharian (Jasanoff, 1987: 109f) and Armenian (Schindler apud Melchert, 1977: 443). A major innovation of CA is the requirement that grammatically neuter nouns functioning as agents must be inflected in a special «ergative» case. For development of the ergative ending **-anti* from the ablative-instrumental of *r/n*-stems see again Garrett, 1990.

Hittite preserves some examples of «apophony» and accent alternation in nominal paradigms: *e.g.* nom.-acc. sg. *tēkan* [tēgan] «earth» *vs.* oblique *takn-* [tagń-]; anim. nom. sg. *āssu-s* «good» *vs.* gen. sg. *āssaw-as* (cf. Skt. *vasú-s vs. vasó-s*). However, such alternations are mostly leveled out in nominal inflection in Anatolian as elsewhere.

Even a summary of nominal derivation in Anatolian is impossible here. For a useful overview see Oettinger, 1986.

3.3. Pronouns

Personal pronouns are predictably irregular in both stem formation and inflection. For the most part the stems are recognizably Indo-European (*e.g.* Hitt. *wēs* «we» *vs. anzās* «us» like Goth. *weis vs. uns*), but some of the inflectional endings have yet to be satisfactorily explained. Note in particular the equation of Hitt. *ammug* «me»), HLuv. *(a)mu*, Lyc. *a/ẽmu* and Lyd. *amu* «I/me» (cf. also Pal. enclitic *-mu* «me»). The initial **em-* of the CA form **emu* is familiar, but the *u*-vocalism of the second syllable (probably borrowed from the second singular accusative form **tu* «thee») is a significant common innovation of the Anatolian languages.

The PIE stem **k*^{w}*i-*, functioning as both interrogative and relative, is pan-Anatolian: Hitt.-Pal.-Luv. *kwi-*, Lyc. *ti-*, Lyd. *qi-*. The demonstrative stem **k̂o/i-* «this» of Hitt. and Pal. *kā/i-* and Luv. *zā/i-* is inherited: cf. Lat. *ci-trā* «on this side», etc. On the other hand, Hitt.-Pal.-Luv. *apā-* «that» (also Lyc. *ebe-* and Lyd. *bi-*, with a secondary meaning «this»)

represents another defining innovation of CA (a virtual PIE **obhó/í-)*. Lydian *αλα-* «other» < PIE **alyo-* is a notable archaism absent elsewhere in Anatolian. The nominative and accusative endings of these pronouns are clearly inherited, but the oblique endings are problematic. The scarcity of evidence outside Hittite makes reconstruction of the CA state of affairs impossible at present.

3.4. Verbal Morphology

It is in the area of verbal morphology where the debate about the relative position of Anatolian referred to in the introduction has raged most fiercely. While many features of verb formation cited below are reassuringly familiar, there are serious and surprising discrepancies between the Anatolian verb and that traditionally reconstructed for PIE. Solutions to this problem are very divergent and mostly irreconcilable. The following summary is as objective and non-committal as possible, but it unavoidably represents a particular point of view, and readers should not assume that it even remotely reflects a consensus.

The Anatolian verb does have the expected three persons. As in the noun, there are only two numbers, singular and plural, but the plural endings we have formally resemble dual endings elsewhere: Hitt. pres. lst pl. *-weni* = CLuv. *-unni*; Hitt. pret. lst pl. *-wen* prob. = Lyd. *-wν* (cf. Skt. lst dual *-va(s)*, etc.).

All languages but Lydian show clear evidence for a middle voice beside the active. Judging from the ample Hittite evidence, we may conclude that the middle marks the familiar functions of passive, reciprocity and actions internal to the subject. Many verbs such as «sit» or «lie» are inflected only in the middle, and Hittite has examples of «deponents» which are middle in form, but active in meaning. The formal marker of the present middle (at least in the third person) is a suffix *-*r*, a feature Anatolian shares with Italic, Celtic and Tocharian: see Yoshida (1990) for details. Other forms of the middle employ a suffix *-*ti*, which is the CA reflexive particle of unknown orogin. Lycian unexpectedly has middle forms with a nasal suffix (*e.g. sijẽni* «lies»), whose source is also problematic (Melchert, 1992a).

The tense-aspect system is surprisingly simple. There is a present tense which also serves for the future, and a preterite. The two tenses are formed from the same stem, being distinguished only by the endings. Present active endings are characterized by the familiar marker *-i*, but the spread of this marker to the middle was incomplete in CA (*e.g.* OH *kitta* «lies» < CA

**k̂eíto* < **k̂eítor*). There are no assured traces of the « augment » prefix **é-* of the PIE imperfect and aorist, nor of the reduplication of the PIE perfect.

There is also no direct evidence for the PIE aspectual distinction of imperfective, perfective and stative marked by the so-called present, aorist and perfect stems (but see further below). The status of aspect in Anatolian remains a topic for research, but it is clear that the suffix **-sk̂e/o-*, usually referred to as « iterative » in Anatolian, is used more generally to indicate imperfectivity.

The are only two moods, indicative and imperative. The second singular imperative usually consists of the bare verbal stem, but there are also traces of the PIE particle**-dhi* (Hitt. *īt* « go ! » = Grk. *íthi*). The imperative second plural expectedly is identical to the preterite indicative, since the latter has no « augment ». There are also third person forms marked with a final *-u* replacing indicative *-i*, comparable to those in Sanskrit (Hitt. *ēstu* = Skt. *ástu* « let it be »).

As to non-finite forms, there is a single participle, generally with a past passive value for transitive verbs, and that of an attained state for intransitives. In Luvian and Lycian the productive form has a suffix **-ómmo-* < **-ómno-*, but there are also lexicalized examples with a suffix **-ónt-*, which is the only form in Hittite. The position of Palaic is unclear, while Lydian goes it own way with a suffix in *-l*. Palaic and Luvian have an infinitive in *-una* < **-unō* (allative of a verbal noun in **-wr̥/-wen*), and Hittite *-wanzi* is some form of the same paradigm. Lycian *-Ṽne* and Hittite *-ānna* similarly continue **-ātnō* (allative of a verbal noun in **-ātr̥/-ātn-*). Once again, Lydian is unique in having a form in - *l*.

Hittite has a famous set of periphrastic forms with the verbs « have » and « be » plus the past participle. Their value appears to be that of an attained state, not that of a perfect as in the modern Western European languages.

Anatolian has representatives of most of the well-established verbal classes of PIE. I cite Hittite evidence for convenience, but most of these types are also attested somewhere else in Anatolian: root presents (*ēszi, asanzi* « be »); « Narten » presents (*wēkzi/wekanzi* « demand »); *nu*-suffix (*tepnu-* « belittle »); nasal-infix (*harni(n)k-* « destroy »); transitive-iteratives with *o*-grade root and suffix **-eye-* (*wasse/a-* « clothe »); deverbatives in **-ye/o-* (*karpiye-* « lift »); denominatives in *-eh_2- (*nēwahh-* « make new ») and extended *-eh_2-*ye/o-* (*parsnāi-* « squat »); denominatives in **-ye/o-* (*tūriye-* « hitch »); statives in *-eh_1- (*marse-* « be false ») and « inchoatives » in *-eh_1-*s-* (*marsess-* « become false »). On the other hand, there are no solid examples of simple thematic stems of the well-known type of **bhére/o-* « carry » (Grk. *phérei*, Stk. *bhárati*). Attempted explanations of this startling fact vary widely, with likewise very different implications for the PIE verb.

We come finally to the greatest conundrum of the Anatolian verb. All the stem types cited above are inflected with endings which are easily derivable from PIE: in Hittite for example the present endings are *-mi*, *-si*, *-zi*, *-weni*, *-teni*, *-anzi* (with regular **-ti* > *-zi* in the third persons). The preterite endings correspond: Hitt. *-un* and Lyd. *-ν* < **-m/-om*, Hitt. *-s*, *-t*, *-wen*, *-ten* and Pal. *-ant(a)* < **-e/ont*. However, beside this «mi-conjugation», Hittite also shows a «*hi*- conjugation», whose present singular endings are *-hhi* (OH *-hhe*), *-ti*, and *-i* (OH *-e*). Ample evidence from HLuvian confirms that this type is CA (Morpurgo Davies, 1979), and there are virtually certain remnants in CLuvian, Palaic, and Lycian.

I follow Jasanoff (1979) in assuming that the principal source of this type is a distinct class of PIE presents with **o/e* vocalism attested only in limited and altered form elsewhere in IE (*e.g.* Hitt. *malla-* = Lat. *molō* «grind»). For very different views see the articles by Cowgill and Kurylowicz in the same anthology, among many others.

The preterite endings of this class in CA are reflected by Pal.- CLuv. *-hha/-ha* = Lyc. *-xa/-ga* for the lst singular, and Hitt. *-ta* and *-s* for the 2nd and 3rd singular respectively. It also seems certain that the generalized Hittite preterite third plural ending *-er* has its origin in this class. One could thus derive this preterite from a conflation of the PIE perfect and sigmatic aorist (an event paralleled elsewhere), but it is by no means clear that this is the proper explanation.

3.5. Syntax

A systematic survey being impossible, I cite only two important innovations of Anatolian. For the development of «split ergativity» already mentioned and the related creation of enclitic subject pronouns with intransitive verbs see Garrett, 1990. The Anatolian languages are also justly famous for complex sequences of enclitic pronouns and «particles» attached to the first word in a sentence: *e.g.* Hitt. DUMU-*ŠU=ma=wa=ssi=za=kan*... or Lyd. *fa=k=mλ=it=in*... . However, differences in both the form and ordering of these elements in the various languages suggest that the attested complexity represents a parallel but independent elaboration of a system which was only embryonic in CA.

3.6. Lexicon

It has often been claimed that Hittite, despite a clearly Indo-European grammatical structure, has largely replaced its inherited vocabulary. This false claim is based on the skewed picture of the Hittite lexicon given us by our very limited evidence. Most of our material deals with the state religion. Since the latter is mostly non-Indo-European, it is hardly surprising that much of the attendant vocabulary is as well. An informal survey using the 100-word list of Swadesh shows that at least 75% of Hittite core vocabulary is inherited or built on inherited material (admitting only the most secure IE etymologies): cf. the similar conclusions of Tischler, 1979: 266f. Our knowledge of the core vocabulary of the other Anatolian languages is too limited to draw firm conclusions, but I know of nothing to suggest that the results would be radically different for them.

4. Conclusion

Although the Anatolian family is among the oldest in Indo-European in terms of the dates of its attestation, it is among the latest to be properly investigated. Difficulties in interpreting the limited evidence of the languages other than Hittite have until recently further impeded the process of integrating Anatolian into the overall picture of Indo-European dialects. The last quarter-century has seen dramatic progress in our understanding of Hittite and the rest of the Anatolian languages, and there is good reason to hope that further new discoveries will allow this progress to continue in the future.

Bibliographie

BADER F. (1985). – « De skr. *anyá-* à skr. *ắrya-* ; noms i.-e. de l'"autre" », *Bulletin de la Société linguistique de Paris* 80, p. 57-90.

EICHNER H. (1973). – « Die Etymologie von heth. *mēhur* », *Münchener Studien zur Sprachwissenschaft* 31, p. 53-107.

– (1986). – « Die Akzentuation des Lydischen », *Sprache* 32, p. 7-21.

GARRETT A. (1990). – « The Origin of NP Split Ergativity », *Language* 66, p. 261-296.

HAWKINS-MORPURGO DAVIES-NEUMANN (1974). – *Hittite Hieroglyphs and Luwian: New evidence for the connection* (= Nachrichten der Akad. Wiss. Göttingen, Phil.-hist. Kl., 1973, no. 6), Göttingen.

JASANOFF J. (1979). – « The position of the *hi*-Conjugation », *in* W. Meid & E. Neu (eds.), *Hethitisch und Indogermanisch*, p. 79-90.

– (1987). – « Some irregular imperatives in Tocharian » in C. Watkins (ed.), *Studies in Memory of Warren Cowgill*, 92-112, Berlin.

KIMBALL S. (1983). – *Hittite Plene Writing*. Ph. D. diss., Univ. of Pennsylvania, Philadelphia.

MELCHERT C. (1977). – *Ablative and Instrumental in Hittite*. Ph. D. diss., Harvard University, Cambridge, Mass.

– (1987). – « PIE Velars in Luvian », in C. Watkins (ed.), *Studies in Memory of Warren Cowgill*, 182-204. Berlin.

– (1989). – « New Luvo-Lycian Isoglosses », *Historische Sprachforschung* 102, p. 23-45.

– (1992a). – « The Middle Voice in Lycian », *Historische Sprachforschung* 105, p. 189-199.

– (1992b). – « Relative Chronology and Anatolian: the Vowel System », *in* R. Beekes *et al.*, *Akten der VIII. Fachtagung der idg. Gesellschaft*, Leiden, 1987, p. 41-53, Innsbruck.

– (1992c). – « Hittite Vocalism », *in* O. Carruba (ed.), *Per una grammatica ittita*, p. 181-196, Pavia.

MORPURGO DAVIES A. (1979). – The Luwian Languages and the Hittite *hi*-conjugation », *in* B. Brogyanyi (ed.), *Festschrift for Oswald Szemerényi*, 577-610, Amsterdam.

– (1982/3). – « Dentals, rhotacism and verbal endings in the Luwian languages », *Zeitschrift für vergleichende Sprachwissenschaft* 96, p. 245-270.

OETTINGER N. (1986). – « *Indo-Hittite* »-*Hypothese und Wortbildung*, Innsbruck.

– (1987). – « Bemerkungen zur anat. *i*-Motion und Genusfrage », *Zeitschrift für vergleichende Sprachwissenschaft*, 100, p. 35-43.

RAY J. (1981). – « An Approach to the Carian Script », *Kadmos* 20, p. 150-162.

TISCHLER J. (1979). – « Der indogermanische Anteil am Wortschatz des Hethitischen », *in* W. Meid & E. Neu (eds.), *Hethitisch und Indogermanisch*, p. 257-267, Innsbruck.

WATKINS C. (1986). – « The Language of the Trojans », *in* M. Mellink (ed.), *Troy and the Trojan War*, Bryn Mawr.

YOSHIDA K. (1990). – *The Hittite Medio-Passive Endings in -ri*, Berlin.

CHAPITRE 7

L'ARMÉNIEN

Charles de LAMBERTERIE

Sommaire. – 1. Histoire de la recherche : (a) avant Hübschmann ; (b) Hübschmann ; (c) de Hübschmann à nos jours : Meillet et Adjarian. – 2. Place de l'arménien parmi les langues indo-européennes : questions d'ensemble. – 3. Phonétique : (a) la « loi de Meillet » ; (b) le traitement des gutturales ; (c) la mutation consonantique : innovation ou archaïsme ? (d) la prothèse vocalique. – 4. Morphologie. – 5. Lexique.

1. HISTOIRE DE LA RECHERCHE

La charnière de cette histoire est constituée par la personnalité du savant allemand Heinrich Hübschmann (1848-1908) [1].

1. Sur les recherches antérieures à Hübschmann, voir Schmitt, 1975 et 1986 ; sur Hübschmann, l'introduction de R. Schmitt à Hübschmann, 1976 (p. IX-XIV). Cf. aussi Zeller, 1927, 85-9.

a) Il est inutile de mentionner les étymologies « sauvages », antérieures au développement de la grammaire comparée, qui consistaient à rapprocher un mot arménien d'un mot grec, latin, hébreu, persan ou turc pourvu que les sens concordent et qu'il y ait une vague ressemblance phonétique. Cette manière de faire est encore pratiquée systématiquement dans le grand dictionnaire de l'arménien ancien, à savoir le *Nor baṙgirkʿ haykazean lezui* publié par les Mékhitaristes de Venise en 1836-37 ; c'est que les auteurs de cet ouvrage – remarquable par ailleurs – étaient restés à l'écart des courants intellectuels de leur époque, ceux notamment qui cherchaient à délimiter des familles de langues.

Si ces travaux n'ont plus maintenant qu'un intérêt historique, il peut arriver que tel ou tel d'entre eux renferme des vues que la postérité a ratifiées et qui font maintenant partie du fonds commun de la linguistique historique. Pour le problème qui nous intéresse, l'arméniste italien G. Bolognesi a montré tout récemment que l'orientaliste allemand J. Schröder avait, dès le début du XVIIIe siècle, relevé plusieurs affinités lexicales entre l'arménien et d'autres langues (arm. *kʿan* = lat. *quam* « que », particule prohibitive *mi* = gr. μή, etc.), et fait le départ entre mots empruntés à l'iranien et mots hérités [2]. Mais Schröder n'avait évidemment pas les moyens de prouver la justesse de ses intuitions, puisqu'à son époque il n'était question ni de régularité des changements phonétiques, ni même de famille des langues indo-européennes, toutes notions qui ne sont apparues qu'au siècle suivant.

La nouveauté apportée par Franz Bopp (1791-1867) a été de substituer aux rapprochements isolés des comparaisons de systèmes flexionnels. Faute, sans doute, de connaître l'arménien, Bopp lui-même n'a pas fait état de cette langue dans son ouvrage de 1816 sur la conjugaison, alors que pourtant la flexion verbale de l'arménien présente avec les autres langues indo-européennes des analogies évidentes : c'est ainsi que le présent de l'indicatif du verbe *berel* « porter » – soit *berem, beres, berē, beremkʿ, berēkʿ, beren* – a, même pour l'œil le moins exercé, une « allure » indo-européenne frappante et se laisse comparer immédiatement à skr. *bhárati*, gr. φέρω, lat. *ferō*, got. *baira*, toutes formes que Bopp citait abondamment et qui servaient de base à ses démonstrations [3]. Quand, en 1833, Bopp publie le premier volume de sa *Vergleichende Grammatik*, l'arménien n'y

2. Schröder, 1711 ; Bolognesi, 1988 et 1990b, avec la conclusion (1990b, 235) : « Il est bon de ne pas négliger *a priori* la lecture de livres appartenant à la "préhistoire" de la linguistique, celle-ci pouvant nous réserver la belle surprise de découvrir qu'en réalité des vérités scientifiques, considérées comme des acquisitions récentes, ont depuis longtemps été énoncées ».

3. Bopp, 1816. – Dans une lettre à Karl Joseph Windischmann, Bopp déclare pourtant qu'il avait parcouru une grammaire de l'arménien au cours de l'été 1814 lors de son séjour à Paris (voir Schmitt, 1975, 4). Il s'agit probablement du bref *Essai sur la langue arménienne* (*Pʿorjaṙutʿiwn haykakan lezuin*) de M. Bellaud, docteur en médecine, publié à Paris en 1812 (an 1261 de l'ère arménienne). Élève de Chahan de Cirbied, qui depuis le 5 février 1811

est toujours pas pris en compte. Mais la parution de ce livre eut un effet séminal en stimulant l'esprit de deux autres savants allemands, qui, à la différence de Bopp, avaient séjourné chez les Mékhitaristes de Venise et pratiquaient l'arménien, à savoir Julius Heinrich Petermann (1801-1876) et Friedrich Windischmann (1811-1861). Ils purent ainsi montrer, le premier dans sa *Grammatica linguæ armeniacæ* de 1837, et le second dans un ouvrage paru en 1846 seulement mais dont l'idée avait germé dès 1833, que l'arménien appartenait à la famille des langues indo-européennes, c'est-à-dire à ce qu'on entendait par là, selon l'enseignement de Bopp, à cette date, à savoir l'indo-iranien, le grec, le latin, le germanique, le baltique et le slave [4]. Bopp lui-même, toujours à l'affût du nouveau et désireux d'élargir le cercle de la famille indo-européenne – ce qui l'entraîna parfois, comme on sait, à des tentatives plus ou moins heureuses –, fut aussitôt convaincu par cette démonstration, et il fait état de l'arménien dans le cinquième volume (1849) de son ouvrage, en se référant aux travaux de Petermann et de Windischmann. Certaines de ses analyses restent vraies aujourd'hui encore, ainsi quand il remarque que pour le vocalisme l'arménien concorde avec le grec et non avec le sanskrit (*ber-* « porter » = gr. φερ- vs skr. *bhar-*), ou qu'il attribue l'*ē* de *berē* « il porte » et de *berēk'* « vous portez », en regard de *-e-* dans le reste du paradigme, au traitement d'un ancien **-t-* [5].

donnait un cours d'arménien à l'École spéciale des langues orientales vivantes, l'auteur de cet opuscule le présente comme une propédeutique à la « grammaire arménienne complète » que préparait son maître (p. VI). Ladite *Grammaire de l'arménien* parut, de fait, en 1823 (voir Dumézil, 1948, 65).

4. Petermann, 1837; Windischmann, 1846 (sur la genèse de cet ouvrage, voir Schmitt, 1975, 11-4). – Le slave, absent du premier volume (1833) de l'ouvrage de Bopp, figure dans le titre à partir du second volume (1835). – Sur l'importance de la *Vergleichende Grammatik*, voir l'introduction de M. Bréal à la traduction française de la 2[e] édition (1866-72, I [1866], XXXVI-XXXVIII). L'admirable texte de Bréal (p. I-LVII), qui porte la marque d'une clairvoyance peu commune, mérite d'être considéré comme un classique de la linguistique.

5. Bopp, 1833-49, 1271-2. L'auteur rapproche aussi, pour la formation du présent, arm. *luanam* de gr. πλύνω « baigner » (1272). À propos des présents en *-a-*, il évoque la première conjugaison latine (1273, avec référence à Windischmann), et sur ce point il a raison plus encore qu'il ne le pense, car l'exemple qu'il cite, à savoir *orsam* « chasser », s'analyse à l'intérieur de l'arménien comme un dérivé de *ors*, gén. *orsoy* « chasse » (sur le type, voir Klingenschmitt, 1982, 89), au même titre qu'en latin *procāre* « faire une demande en mariage » dérive de *procus*, gén. *procī* « prétendant ». À partir de la racine i.-e. **perk̂-/prek̂-* « chercher à obtenir », les deux langues présentent un remarquable parallélisme de développement : pré-arménien **pórk̂-* $^{e}/_{o}$- (nom d'action) → **pork̂-āy* $^{e}/_{o}$-, pré-latin **prok̂-ó-* (nom d'agent) → **prok̂-āy* $^{e}/_{o}$-. L'arménien atteste en outre un thème nominal **pork̂-ā-* dans le nom d'agent composé *jkn-ors*, instr. °*orsaw* « pêcheur », *litt.* « chasseur de poisson » (sur le type, voir Meillet, 1962, 171-4 = *MSL* 18, 1913, 257-60).

La première mention de l'arménien apparaît pp. 1269-70, à propos de l'infinitif en *-l* que Bopp rapproche de l'infinitif germanique en *-an* et des formations indo-iraniennes en *-ana-*, en considérant que *-l* provient de la même « dégénérescence » de *-n* que dans arm.

La linguistique comparative du milieu du XIX[e] siècle, dominée par la personnalité de Bopp, partage, pour l'essentiel, les qualités et les défauts de cet auteur. Ce qui fait la force de Bopp, c'est, outre une curiosité d'esprit peu commune, que sa pratique de la comparaison se fonde moins sur le lexique que sur la morphologie et qu'il propose une explication des flexions [6]. Sa faiblesse principale, c'est l'absence de toute rigueur en matière de phonétique historique, qui le conduit la plupart du temps à des rapprochements inadmissibles parce que contradictoires : tel phonème, tel morphème de la langue-mère est en général représenté dans une langue indo-européenne par tel phonème ou tel morphème, mais « peut » aussi, et sans raison apparente, aboutir à tel autre. Dans le cas de l'arménien, langue qui a beaucoup emprunté à l'iranien, ce laxisme empêchait évidemment de faire le départ entre mots hérités et mots empruntés, si bien que l'arménien apparaissait comme une langue iranienne. Telle est la doctrine qui, à partir de l'ouvrage de F. Windischmann, a prévalu jusqu'à Hübschmann et qui est enseignée notamment dans les deux grandes synthèses antérieures au courant néo-grammairien, à savoir le *Compendium* d'A. Schleicher [7] et la deuxième édition de la *Vergleichende Grammatik* de Bopp. Comme s'il avait mauvaise conscience d'avoir presque entièrement négligé l'arménien dans la première édition, Bopp consacre à cette langue le début de l'introduction de la deuxième, avec le projet de « mettre en lumière les rapports, quelquefois très-cachés, et en partie encore inconnus, qui l'unissent au sanscrit, au zend et aux idiomes congénères de l'Europe » [8]. Admettant que « l'arménien fait partie du rameau iranien de notre famille de langues » [9], Bopp constate qu'il semble cependant s'en écarter sur certains points et attribue ces traits aberrants au fait que la forme d'iranien dont procède l'arménien serait un état de langue plus ancien que l'iranien ancien. Il met le doigt sur certaines difficultés, et signale notamment qu'en arménien ancien « la racine sanscrite *dā-* "donner" était devenue *t* (*tam* "je donne") *d'après une loi de substitution analogue à celle des langues germaniques* », puis « a repris dans la prononciation actuelle [*sc.* occidentale] la valeur de *d* : de sorte que nous avons aujourd'hui une forme

ayl, gr. ἄλλος, lat. *alius* « autre » en regard de skr. *anyá-* de même sens. Dans la lignée de Bopp, divers auteurs ont cherché à ramener à l'unité les deux noms indo-européens **alyo-* et **anyo-* de l'« autre » (ainsi, récemment, Bonfante, 1978, qui part de **alyo-*), tandis que d'autres les considèrent comme irréductibles l'un à l'autre (ainsi Bader, 1985, 57), ce qui paraît préférable : l'élément **-n-* qui figure dans **anyo-* ne saurait, en effet, être séparé du thème pronominal i.-e. **no-* qui désigne l'objet lointain (v.sl. *onŭ*, etc.), et ce dernier est représenté en arménien par *n-* (anaphorique *na* < **no-ay*, démonstratif *ayn* < **ay-no-*, etc.).

6. Voir sur ce point Bréal, *op. cit.* (n. 4), XIX-XXII (à propos de Bopp, 1816).

7. Schleicher, 1861-62, 5, et même doctrine déjà dans un ouvrage de 1850, à la suite de Windischmann (voir Schmitt, 1975, 14).

8. Bopp, 1866-72, I (1866), 11 (= 1857-61, I, 1857, XIV).

9. *Ibid.*, 15 (= XVII).

dam "je donne" qui répond au sanscrit *dadāmi*, et *das* "tu donnes", qui sonne comme la forme équivalente en latin »[10]. Mais dans d'autres mots, c'est au contraire l'arménien moderne qui « ressemble au germanique » et « l'ancienne prononciation » qui « se rapproche davantage du sanscrit »[11]. Tout Bopp est là, avec ses contradictions, son incapacité à maîtriser la phonétique historique et à dépasser le stade de la ressemblance, mais aussi avec une intuition féconde en matière d'analyse morphologique[12].

Au point où nous en sommes, l'arménien apparaît comme un idiome qui, la plupart du temps, concorde avec l'« aryen » (= indo-iranien), si bien qu'on peut le considérer comme une langue iranienne, mais qui, sur plusieurs points importants, s'en sépare. Pour rendre compte de ces contradictions, on recourait le plus souvent au concept, bien commode, de « langue mixte ». C'est ainsi que procède notamment l'un des personnages les plus influents de cette époque, à savoir Paul de Lagarde (Bötticher de son nom de naissance, 1827-1891), qui distingue en arménien trois strates, toutes les trois iraniennes : la couche « haïkanienne » (du nom de *Hayk*, l'ancêtre mythique du peuple arménien), la plus ancienne, puis la couche arsacide, remontant à l'époque de la domination parthe en Arménie, et enfin la couche sassanide, constituée de mots perses[13]. L'idée d'établir une chronologie relative dans le lexique arménien est en elle-même un progrès notable, mais le point de départ reste toujours l'iranien : là où l'haïkanien s'en écarte (système vocalique plus proche du grec que de l'indo-iranien, mutation consonantique), c'est, croit-on, en vertu d'une évolution particulière.

---

10. Bopp, 1866-72, I (1866), § 183^{b}, 2, p. 402 (« *d'après... germaniques* » souligné par nous).

11. *Ibid.*, § 87, 1, p. 149.

12. Cette analyse de arm. *ta-* = lat. *dă-* < i.-e. **də-* a été reprise notamment par R. Godel (1982, 21 = *REArm N.S.* 2, 1965, 23), mais avec des réserves, car le présent arménien peut aussi reposer sur **də-y* $^{e}/_{o}$- (Meillet, 1936, 132-3). G. Klingenschmitt (1982, 85-6) estime que le seul présent de la racine **dō-* < **deh*$_3$- que l'on puisse poser en indo-européen est une formation à redoublement, reconstruite à partir du grec et de l'indo-iranien, et laisse la forme arménienne inexpliquée ; mais la racine hittite *dā-* « prendre », que l'on s'accorde à faire dériver de i.-e. **dō-*, a un présent sans redoublement. Tout n'est certes pas clair dans la flexion du verbe arménien, mais le thème *tac'-* du subjonctif aoriste, qui repose sur i.-e. **də-sk̂* $^{e}/_{o}$-, se retrouve dans l'itératif hittite *daskizzi* de *dā-* (Watkins, 1969, 57) ; on rapprochera aussi, avec désinences secondaires, l'itératif ionien δόσκον. La question reste posée de savoir si ce subjonctif, qui en synchronie se rattache à l'aoriste, doit être considéré en diachronie comme un dérivé du thème d'aoriste ou du thème de présent ; R. Godel a donné successivement les deux analyses (1982, 34 = *REArm* 2, 1965, 36 ; 1982, 85-6 = *REArm* 14, 1980, 16).

13. Voir le détail des données chez Schmitt, 1975, 15-9.

b) Le grand mérite du courant néo-grammairien, qui s'est développé en Allemagne à partir de 1875, a été d'introduire une exigence de rigueur en matière de phonétique historique. L'idée qu'on se faisait jusqu'alors du système phonétique de l'indo-européen s'est trouvée bouleversée dès lors qu'on a posé en principe le primat de la correspondance sur la ressemblance. La rupture avec les méthodes de Bopp et de Schleicher étant consommée, il en est résulté pour la reconstruction de l'indo-européen une série d'acquis durables, sur lesquels nous vivons encore même si les générations suivantes ont été amenées à les modifier sur certains points : théorie des gutturales, des sonantes voyelles (brèves et longues), du vocalisme et de l'alternance vocalique. Au nom du principe de la « constance des lois phonétiques », il est devenu nécessaire de rendre compte des anomalies au lieu d'admettre, comme on le faisait antérieurement, la possibilité de déviations arbitraires. Parmi les résultats les plus importants de cette démarche, il faut mentionner deux grands articles publiés dans le XXIIIe volume de la *Kuhns Zeitschrift*. L'un, du danois Karl Verner (1846-1896), explique une exception à la mutation consonantique du germanique en établissant ce qu'on appelle depuis lors la « loi de Verner », à savoir une trace des oppositions accentuelles indo-européennes en germanique ancien [14]. L'autre, de H. Hübschmann, traite des contradictions phonétiques de l'arménien. Ainsi le nom indo-européen *g^wow-* du « bovin » est représenté par *kov* (gén. *kovu*) « vache », et la racine **aĝ-* « mener » par *acem* « je mène » ; mais on trouve d'autre part *gaw-* et *-az-* dans le composé *gawazan* « houlette ». L'auteur montre qu'il faut distinguer dans le lexique arménien ce qui est emprunté à l'iranien (*gaw-az-*, cf. av. *gavāza-*, pers. *gavāz* « aiguillon ») de ce qui est hérité (*kov, ac-)*, et qu'une fois mis de côté les emprunts, qui sont très nombreux, il subsiste en arménien un élément irréductible à l'indo-iranien. L'arménien doit donc être considéré comme une branche autonome à l'intérieur de la famille des langues indo-européennes [15].

Par la rigueur de sa méthode, Hübschmann est le véritable fondateur de la grammaire comparée de l'arménien ; même si l'on trouve des idées justes chez tel ou tel de ses devanciers, proches ou lointains, il est le premier qui ait établi sur des bases solides l'étude préhistorique de la

---

14. Verner, 1877. Tout récemment, B. Olsen, compatriote de Verner, a pensé pouvoir établir un équivalent arménien de cette loi (1989) ; l'idée est en elle-même fort intéressante, mais les exemples allégués ne nous paraissent pas tous également convaincants.

15. H. Hübschmann, « Ueber die stellung des armenischen im kreise der indogermanischen sprachen », *KZ* 23, 1 (1875), 5-49, repris dans Hübschmann, 1976, 1-45. Les exemples cités ici sont analysés p. 9, 20, 21 et 33 (= 5, 16, 17, 29 du recueil). La finale de arm. *gawazan* (instr. sg. *gawazanaw*) est un suffixe *-an-* usuel dans la langue pour former des noms d'instruments (voir ci-dessous § 5).

langue arménienne [16]. L'essentiel des résultats auxquels il est parvenu est rassemblé dans un dictionnaire étymologique de l'arménien, première partie d'une *Armenische Grammatik* dont la suite n'a jamais paru du fait de la mort de l'auteur; le lexique arménien y est divisé en deux grands groupes : mots empruntés et mots hérités (« *echtarmenische Wörter* ») [17].

À la fois iraniste et arméniste, Hübschmann était en son temps le mieux placé pour identifier les emprunts de l'arménien à l'iranien, et il a été l'initiateur d'un champ de recherche qui après lui n'a cessé de se développer. L'arménien classique est la langue dans laquelle est écrit un corpus de textes dont la couche la plus ancienne, à savoir une traduction de la Bible chrétienne, date du Ve siècle de notre ère. Comme la seule langue iranienne à peu près contemporaine que l'on connaissait au XIXe siècle était le pehlevi ou moyen perse, intermédiaire entre le vieux perse et le persan, Hübschmann qualifiait de « *persische Wörter* » les originaux iraniens des emprunts arméniens; quand il ne disposait pas d'une forme pehlevie, il remontait à l'iranien ancien (avestique et vieux perse) ou descendait jusqu'au persan. Mais par la suite les perspectives ont changé. Grâce à la découverte des textes de Turfan et à l'enrichissement du corpus des inscriptions, nous savons désormais que la plupart des mots empruntés par l'arménien classique à l'iranien proviennent du parthe, dialecte iranien du nord-ouest, et remontent à un état ancien de cette langue [18]. C'est en effet à partir du Ier siècle de notre ère que les mots parthes sont entrés en masse dans la langue arménienne, quand le pays a été dominé par une aristocratie arsacide; les premiers monuments de la langue parthe, que nous connaissons par les fouilles de Nisa et par

---

16. Voir sur ce point Meillet, 1936, 15 et 1937, 476. – En ce qui concerne le « mouvement de 1875 », l'exposé de Meillet (1937, 468-78 – cette expression p. 478) garde encore toute sa valeur. Il est fort surprenant que dans un ouvrage de 275 pages consacré aux néo-grammairiens (Jankowsky, 1972) le nom de Hübschmann n'apparaisse pas une seule fois, autant du moins qu'on puisse en juger d'après l'index placé à la fin du volume. Peu importe, en l'occurrence, que Hübschmann ait fait partie ou non du groupe de Leipzig; l'essentiel n'est pas l'appartenance à une chapelle ou le rattachement à un slogan. La fécondité du courant néo-grammairien est d'avoir opéré une rupture épistémologique avec le romantisme (« scission... définitive », Meillet, 1937, 468-9) et introduit une exigence de rigueur en linguistique historique; à cet égard, Hübschmann s'y rattache pleinement et son œuvre en est l'un des acquis les plus brillants. Il en va de même du *Mémoire* de F. de Saussure (1878), quelles qu'aient été les relations de Saussure avec « ces messieurs de Leipzig ».

17. Hübschmann, 1897. La plupart des études rassemblées dans Hübschmann, 1976 étant antérieures à cet ouvrage, on peut considérer que, pour l'essentiel, il représente le dernier point de vue de l'auteur.

18. Ce fait a été établi par A. Meillet, *MSL* 17, 1911, 242-50 (repris dans Meillet, 1977, 142-50); *REArm* 1, 1920, 9-14 (= 1977, 180-85) et 2, 1922, 1-6 (= 1977, 197-202). De nombreux emprunts à l'iranien ont été identifiés aussi par É. Benveniste (cf. Bolognesi, 1984) et par bien d'autres savants : voir les synthèses de Bolognesi, 1960; Schmitt, 1983 et 1987.

quelques rares inscriptions, datent, pour l'essentiel, du 1er siècle av. J.-C. et se distinguent des documents postérieurs (c'est-à-dire les inscriptions d'époque sassanide et les textes manichéens de Turfan) par certains traits, comme la conservation des occlusives sourdes intervocaliques, que l'on retrouve dans les emprunts arméniens.

Les principes posés par Hübschmann demeurent, aujourd'hui encore, plus valables que jamais : le moindre soupçon qu'un mot arménien soit emprunté à l'iranien mérite d'être pris en considération[19]. Nous sommes même en droit de considérer comme emprunté plus d'un mot que l'on croyait jadis hérité, même si nous ne disposons pas de l'original iranien; comme il n'existe aucune langue moyen-iranienne dont le lexique conservé soit aussi étendu que celui de l'arménien classique, l'arménien est parfois le seul vestige d'un mot iranien disparu ailleurs. C'est ainsi que le correspondant iranien **niδrā-* de skr. *nidrā-* «sommeil» n'est connu que par l'emprunt arménien *nihr*[20]. À cela s'ajoute que l'influence de l'iranien sur l'arménien ne se limite pas aux emprunts de mots; elle s'étend parfois aux emprunts de sens, sous la forme de calques. Un exemple caractéristique à cet égard est celui du verbe *ənt'eṙnum*, aor. *ənt'erc'ay* «lire», qui repose sur un composé **əndi-harc'-* de *harc'-* «interroger» (prés. *harc'anem*, aor. *harc'i*); tant la racine *harc'-* (< i.-e. **pr̥sk̑-*) que le préverbe **əndi-* (< i.-e. **anti*) sont hérités, mais le procédé qui consiste à former un verbe «lire» en ajoutant un préverbe à un verbe «interroger» est emprunté à l'iranien, à preuve v.p. *pars-* «examiner» (le correspondant étymologique de arménien *harc'-*) et *pati-pars-*, pth. **pdfwrs-** */padfurs-/* «lire»[21].

D'une manière générale, la langue arménienne a montré une grande faculté d'assimilation à l'égard des éléments étrangers. En plus des emprunts à l'iranien, elle a beaucoup pris au syriaque et au grec, parfois d'ailleurs au grec par l'intermédiaire du sémitique. L'influence culturelle du grec a été considérable en Arménie au cours de l'époque classique, notamment (mais non uniquement) par le biais de l'école dite hellénistique. Ici aussi, on relève nombre de calques, et bien des mots arméniens sont incompréhensibles sans un détour par le grec. Ainsi, à partir de *p'aṙ-k'* «gloire» (< pth. **frh** */farrah/*), qui traduit gr. δόξα, on obtient le composé *uł-a-p'aṙ*

---

19. Cela est vrai particulièrement lorsque l'arménien et l'iranien connaissent le même développement phonétique à partir de l'indo-européen : les deux langues représentent en effet i.e. **bh* et **dh* par *b* et *d*, etc. Lorsque les critères phonétiques ne permettent pas de trancher, il faut recourir à la morphologie pour décider entre l'emprunt et l'héritage (voir certains cas de ce type Lamberterie, 1984, 236). Parfois l'arménien offre un mélange inextricable de formes empruntées et de formes héritées (Lamberterie, 1985 et 1989).

20. Voir en dernier lieu Bolognesi, 1990a, 389-90 (repris de *REArm* 18, 1984, 294).

21. Cette remarquable étymologie, qui est due à O. Szemerényi (1991, 1888-90 = *Sprache* 12, 1967, 223-5), nous paraît résister aux objections de Klingenschmitt, 1982, 251-2. – Pour d'autres exemples de calques, voir, outre les études citées n. 18, Bolognesi, 1990a, 197-212 (= *RIL* 96, 1962, 235 sqq.) et 319-46 (= *HA* 75, 1961, 657-84).

«ὀρθόδοξος» (de *uğiğ* «droit»), le second membre du composé arménien recevant, sur le modèle de gr. °δοξος, le sens de «doctrine» qu'il n'a aucunement par ailleurs. Le slave a recouru, lui aussi, au même procédé, d'où le composé r. **православный** «orthodoxe», formé sur **слава** «gloire» avec le même suffixe que le simple **славный** «glorieux».

c) L'œuvre que Hübschmann n'a pu écrire, à savoir une grammaire de l'arménien conçue selon les principes de la linguistique historique, a été réalisée par le linguiste français Antoine Meillet (1866-1936), à qui l'on doit non seulement une grammaire comparée de l'arménien et une description strictement synchronique de la langue classique, mais encore une infinité d'études de détail, dont plusieurs sont de véritables livres et dont aucune n'est indifférente [22]. En matière de grammaire comparée de l'arménien, l'œuvre de Meillet constitue, aujourd'hui encore, le point de référence et la base de tous les travaux ultérieurs [23]. Le complément indispensable en est le dictionnaire étymologique écrit entre 1926 et 1935 par le linguiste arménien H. Adjarian (1876-1953), ouvrage admirable tant par la richesse des données qu'il contient que par sa doctrine [24] : contemporain du *Dictionnaire étymologique de la langue latine* d'Ernout et Meillet, il est conçu selon les mêmes principes, et l'on y trouve en outre une bonne part de l'enseignement oral de Meillet.

2. QUESTIONS D'ENSEMBLE

Du fait de la masse des emprunts, la composante héritée de l'indo-européen ne représente dans l'ensemble du lexique arménien qu'une part assez modeste. Le comparatiste travaille donc sur un matériel assez réduit. Si l'on ajoute à cela que la langue n'est attestée qu'à une date tardive

22. Ouvrages : Meillet, 1936 et 1913. Articles : Meillet, 1962 (cinq grands articles, écrits entre 1897 et 1911, qui forment une suite intitulée par l'auteur «Recherches sur la syntaxe comparée de l'arménien», plus une étude sur la composition) et 1977. Le recueil de 1977, dont la réalisation laisse à désirer, est loin d'être complet : voir Lamberterie, 1988 (avec bibl.). Sur l'œuvre de Meillet, voir aussi Bolognesi, 1987.

23. Parmi ces derniers, on peut signaler, pour se limiter aux ouvrages d'ensemble : Godel, 1975; Schmitt, 1981; Jahukyan, 1982 et 1987. – Les travaux publiés entre 1936 (année de la mort de Meillet) et 1972 sont recensés par Schmitt, 1972 [74]; après cette date, il faut se référer à la rubrique «arménien» de la *Bibliographie linguistique* et de l'Indogermanische Chronik (R. Schmitt) dans la revue *Die Sprache*.

24. Adjarian, 1971-79.

(Ve siècle de notre ère) et qu'elle a subi au cours de sa préhistoire de profonds bouleversements, on conçoit qu'il ne soit guère aisé de définir avec précision la place de l'arménien parmi les langues indo-européennes. Certains linguistes dressent des listes de correspondances (empruntées souvent à des sources vieillies) entre l'arménien et telle ou telle langue, et en tirent des données chiffrées qui sont censées révéler des degrés plus ou moins étroits de parenté ; une telle méthode, déjà contestable en elle-même, l'est encore plus dans le cas de l'arménien.

Un point toutefois se dégage assez nettement : il semble bien que les langues avec lesquelles l'arménien présente les isoglosses les plus notables et les plus nombreuses soient le grec et l'indo-iranien ; on relève aussi des correspondances précises avec le phrygien et l'albanais, mais le témoignage qu'on peut tirer de ces deux langues pour la comparaison est fort limité. Tel est le résultat auquel aboutissent la plupart des travaux qui traitent de la question [25]. Grec, arménien, phrygien et indo-iranien constituent notamment ce que l'on appelle « l'aire de l'augment » : skr. *ádhāt* = gr. dial. (ἀν)έθη = arm. *ed* « il plaça » < **édhēt*, aoriste radical dont on rapprochera les formations suffixées v. phryg. *edaes* « fecit », gr. ἔθηκε ; skr. *ábharat* = gr. ἔφερε « il portait » = arm. *eber* « il porta » < **ébheret*. Cette aire coïncide avec celle de l'aoriste thématique : gr. ἔϜιδε « il vit (he saw) » = arm. *egit* = skr. *ávidat* « il trouva » < **éwidet* ; on a même pu montrer qu'en grec homérique l'aoriste Ϝίδε avait conservé dans une formule archaïque le sens de « recevoir, obtenir » qu'il a en arménien et en indo-iranien [26]. Depuis une étude classique de C. Watkins, on sait que l'augment est une ancienne particule de phrase qui fonctionne encore comme telle dans le louvite *a* « et, alors » [27] ; la valeur spécifique qu'a cet élément **é*- dans une partie du domaine indo-européen est donc une innovation commune à ces dialectes et non, comme on le croyait jadis, une conservation de l'état indo-européen originel ; l'aire dialectale ainsi définie n'en est que mieux marquée [28]. L'aoriste thématique apparaît, lui aussi, comme une catégorie récente en indo-européen [29].

---

25. Voir en ce sens notamment l'ouvrage classique de Porzig, 1954, 152 sq. (avec renvoi aux travaux antérieurs, dont le plus important est Meillet, 1922). Parmi les études plus récentes : Solta, 1960 ; Jahukyan, 1980 ; 1982, 199-205 ; 1987, 86-222 ; Schmidt, 1980 et 1990 ; Kortlandt, 1986 ; Lamberterie, 1986. – En ce qui concerne les relations entre arménien et phrygien, les travaux de Bonfante (1946, 1982) et de Haas (1966, 240-57) sont vieillis sur le plan philologique ; les données paléo-phrygiennes sont à citer désormais d'après Brixhe-Lejeune, 1984, ouvrage qui se refuse à la comparaison mais grâce auquel on dispose enfin d'une base solide pour tout futur travail comparatif (cf. aussi Brixhe, 1983).

26. Voir Lamberterie, 1978-79.

27. Watkins, 1963, 15, avec les compléments de Bader, 1973, 32-3.

28. Voir Schmidt, 1980, 39 et n. 13 (avec bibl.).

29. Voir Watkins, 1969, § 43 (avec bibl.).

À l'intérieur de cet ensemble, il est établi, depuis les travaux de H. Pedersen – un des grands noms de la linguistique arménienne (1867-1953) –, que l'arménien est particulièrement proche du grec [30], au point que le comparatiste américain E.P. Hamp a pu parler d'une unité dialectale originelle à laquelle il donne le nom d'« helléno-arménien », comparable dans une certaine mesure au balto-slave [31]. Il est peut-être hardi d'aller jusque-là, encore que l'hypothèse dispose de bons arguments; ce qu'il y a de sûr, en tout cas, c'est que l'arménien, malgré la date tardive à laquelle il est attesté, apparaît comme une sorte de trait d'union entre le grec et l'indo-iranien.

Une tradition ancienne, qui remonte à Hérodote (VII, 73), veut que l'Arménie ait été une colonie phrygienne ('Αρμένιοι, ... ἐόντες Φρυγῶν ἄποικοι). La linguistique n'y contredit pas, et l'on pourrait même songer à rapprocher le fameux mot phrygien βέκος « pain » – mot de la langue la plus ancienne de l'humanité selon Hérodote (II, 2) – du verbe arménien *bekanem*, *beki* « rompre » et du dérivé nominal *bekor* « morceau » (pour le sens, cf. gr. mod. ψωμί « pain » < gr. ancien ψωμίον « petit morceau, bouchée »). Mais nous connaissons trop peu le phrygien pour aboutir à des conclusions fermes [32] ; est-il bien sûr d'ailleurs que le phrygien d'Hérodote soit le même que celui des inscriptions paléo- et néo-phrygiennes ? L'origine du peuple arménien demeure mystérieuse, et l'on a trop souvent sollicité à l'excès les données linguistiques pour résoudre les problèmes d'ethnogénèse. À plus forte raison la réserve s'impose-t-elle à l'égard des tentatives qui cherchent à établir des relations entre l'arménien et des langues anciennes du Proche-Orient dont nous ignorons presque tout.

Il est probable qu'il existait des dialectes dans l'Arménie ancienne, mais la langue littéraire qu'est l'arménien classique obéit à une norme rigoureusement fixée; c'est ainsi qu'elle a pu, à partir de l'invention de l'alphabet et de la traduction des Livres saints par *Mašt'oc'* (*Mesrovb*) et son école dans la première moitié du Ve siècle, être le ciment de l'unité culturelle de la nation. Certains linguistes ont cherché à dresser des listes de doublets étymologiques ou de traitements phonétiques divergents qui témoigneraient de variations dialectales [33] ; mais le matériel allégué ne permet guère d'aboutir à des conclusions décisives. Certains exemples paraissent pourtant incontestables : ainsi le substantif *apa-t'arc'* « apostrophe », composé forgé par l'école hellénistique sur le modèle de gr. ἀπόστροφος, atteste une variante *t'arc'*- (avec deux sourdes aspirées) de la racine *darj*-

---

30. Pedersen, 1982, 220 (= *ZVS* 39, 1906, 442) et 302-9 (écrit en 1924); Meillet, 1936, 142; Bonfante, 1937 et 1982.

31. Voir notamment Hamp, 1979, 4-5 et 1983, 6.

32. D'où le scepticisme de Pedersen, 1925, 44-51.

33. Voir en ce sens Winter, 1966; Jahukyan, 1983 (= 1984), et la préface de Jahukyan à Greppin-Khachaturian, 1986, 9-10.

«tourner» (< i.-e. **dhr̥ǵh-*, cf. gr. τρέχω «tourner sur un axe») qui dans les textes classiques traduit gr. στρέφω [34]. Il est possible aussi que certains dialectes modernes continuent des dialectes anciens différents de la langue classique et soient dignes d'être pris en compte pour la reconstruction du proto-arménien; mais les résultats obtenus actuellement sont encore modestes et fragiles, car les données sont complexes et souvent peu fiables. Si les recherches entreprises en ce domaine sont conduites avec méthode, comme c'est le cas de celles de J. Weitenberg, on peut en attendre beaucoup [35]. On est bien parvenu, en appliquant avec rigueur la méthode comparative, à identifier, sur la base de langues du Nouristan qui ne sont connues qu'à l'époque moderne, une branche de l'indo-iranien qui ne se confond ni avec l'indo-aryen, ni avec l'iranien [36]; pourquoi ne pourrait-il en être de même pour l'arménien? Sans remonter aussi haut, rien n'interdit *a priori* de croire que certains dialectes arméniens modernes aient conservé des archaïsmes comparables à ceux du tsaconien en regard des autres dialectes du grec moderne. Mais les risques d'anachronisme sont évidents : plus d'un romaniste serait tenté, si nous ne connaissions pas le latin, de faire remonter le verbe français *tenir* à un étymon **tenīre* comparable au grec τείνω (< *τέν-yω) «tendre», c'est-à-dire de le situer au même niveau chronologique que esp. *tener* < lat. *tenēre*.

3. PHONÉTIQUE

On n'envisagera ici que quelques points particuliers, instructifs pour la méthode.

a) La «loi de Meillet»

Le groupe de consonnes i.-e. **dw-* (**-dw-*) a subi entre l'indo-européen et l'arménien une évolution remarquable, connue sous le nom de «loi de Meillet», qui montre assez la profondeur des bouleversements intervenus

34. Sur cette racine et ses correspondants grecs, voir Létoublon-Lamberterie, 1980.

35. Voir, par exemple, Weitenberg, 1985 et 1986 (avec bibl.); cf. aussi Jahukyan, 1985. – Sur les dialectes arméniens, voir Adjarian, 1909 et Greppin – Khachaturian, 1986.

36. Voir la contribution de M. Mayrhofer dans le présent recueil.

dans la préhistoire de l'arménien. L'exemple canonique est celui du numéral « deux », i.-e. _*dwō_ > arm. _erku_ (cf. aussi, en composition, _*dwi-_ > _erki-_), que citait volontiers Meillet pour illustrer le primat de la correspondance sur la ressemblance en matière de parenté de langues[37]. Dans une étude récente, nous avons tenté de restituer les étapes suivantes[38] : _*dw-_ > _*dg_^{w}_-_ (renforcement du _*-w-_ appuyé) > _ðg_^{w}_-_ (spirantisation de l'initiale en position implosive) > _*rg_^{w}_-_ (fait banal en phonétique générale, cf. ir. _-ð-_ > arm. _-r-_ entre voyelles dans les emprunts) > _erg_^{w}_-_ (prothèse vocalique, l'arménien n'admettant pas d'initiale _*r-_) > _erk-_ (mutation consonantique, cf. _*g_^{w}_-_ > _k-_ dans i.-e. _*g_^{w}_enā̆_ > _kin_ « femme »). Le traitement est en partie parallèle à celui du groupe i.-e. _*tw-_ > arm. _k'-_ (_k'o_ « ton » < i.-e. _*twos_), à travers les étapes _*tk_^{w}_-_ > _*þk_^{w}_-_ > _*k_^{w}_-_ > _k'-_ (cf. _lik'-_ « laisser » < i.-e. _*lik_^{w}_-_) ; la seule différence est que la spirante sourde _*þ_ s'amuït devant consonne, comme dans le groupe _*tr-_ > _*þr-_ > _*r-_ > _er-_ (_*treyes_ > _erek'_ « trois »), au lieu de dégager un _*r_ comme le fait la sonore _*ð_. Nous pouvons ajouter maintenant qu'une confirmation de cette chronologie est fournie par l'adjectif _merk_ « nu », si l'on admet, avec É. Benveniste, qu'il repose sur i.-e. _*meg_^{w}_-do-_ : il suffit de considérer que dans ce mot la première étape a été une métathèse – phénomène courant dans la préhistoire de l'arménien – aboutissant à _*medg_^{w}_o-_, après quoi l'évolution est identique à celle du numéral « deux »[39].

Ce traitement, qui caractérise en propre l'arménien, paraît singulier à tous égards ; il est cependant conforme à cette tendance générale de la langue qu'est l'élimination des groupes de consonnes tautosyllabiques, traités par disparition du premier élément (_*k͡l-_ > _*ł-_ > _l-_) ou interversion (_*dr-_ > _*ard-_ > _art-_, etc.).

---

37. Ainsi Meillet, 1925, 5-7, 31-4, 106-9. Sur cette correspondance, qui a été souvent – et bien à tort – contestée, voir Lamberterie, 1988, 222-8, avec histoire de la question et relevé de l'ensemble des exemples ; en dernier lieu Vennemann, 1986 (pour la loi), Kortlandt, 1989 (contre la loi) et Picard, 1989 (pour la loi, mais contre les vues de Vennemann). Les adversaires de la loi posent une correspondance i.-e. _*dw-_ > arménien _k-_ (parallèle à _*tw-_ > _k'-_) qu'ils sont bien en peine de justifier par une étymologie convaincante.

38. Lamberterie, 1988, 227.

39. Benveniste, 1930, mais avec une chronologie différente de la nôtre : l'auteur situe la métathèse après la mutation consonantique, alors que nous y voyons une étape antérieure, comme dans le groupe _*dr-_ > _*ard-_ > _art-_.

b) Le traitement des gutturales

Les comparatistes du XIXe siècle divisaient les langues indo-européennes en deux ensembles dialectaux bien différenciés : un groupe occidental ou *centum*, caractérisé par une corrélation *k^w/k, et un groupe oriental ou *satəm* (d'après le nom du numéral «cent» en avestique), caractérisé par une palatalisation des dorsales, d'où une corrélation *$k/\widehat{k}$. Mais la découverte du tokharien et le déchiffrement du hittite ont bouleversé les perspectives : étant donné que ces langues, quoique situées à l'est du domaine, ont conservé la corrélation k^w/k de l'indo-européen commun, on ne peut plus parler d'un groupe *centum*. En revanche, comme il y a une contiguïté géographique des langues *satəm* (indo-iranien, arménien, balto-slave – le cas du phrygien est discuté), on peut encore parler d'une aire *satəm*.

L'arménien occupe ici une situation charnière. D'une part, en effet, il présente, comme les autres langues *satəm*, les corrélations $k/\widehat{k}$, *$g/\widehat{g}$ et *$gh/\widehat{g}h$- [40] ; mais d'autre part on trouve dans quelques mots une trace indirecte des labiovélaires, sous la forme d'un *w*-umlaut : ainsi la racine i.-e. *h_3eng^w- «oindre» (lat. *ungu-*, skr. *añj-*) est-elle représentée par *awc-*, qui repose sur un degré zéro *$h_3\mathring{n}\, g^w$- à travers les étapes *ang^w- > *awng- (umlaut) > *$aw\widehat{g}$- (écrasement de la nasale, et palatalisation de *g au contact de *-w-*) > *awc-* (mutation consonantique) [41].

40. Sur le détail des données, voir notamment Jahukyan, 1978, 119-29 (repris de *ZVS* 89, 1975, 31-42 [version allemande]).

41. Voir en dernier lieu Bolognesi, 1990a, 371-5 (avec bibl.); Lamberterie, 1990, 268-70, où il est montré que l'anticipation de l'élément labial se retrouve en grec dans le nom du «cou», gr. αὐχήν, arm. *awjik'*, qui est un dérivé à suffixe *-*w*- de la racine *$an\widehat{g}h$- «resserrer» (gr. ἀγχ-, arménien *anj-*).
Nous ne poserons pas ici le «problème des gutturales», qui consiste à savoir s'il faut restituer en indo-européen deux séries de dorsales, sc. *k^w/k, conservées dans les langues *centum* et palatalisées en *$k/\widehat{k}$ dans l'aire *satəm*, ou trois séries, sc. *k^w (labiovélaire), *k (vélaire pure, représentée par *k* dans l'ensemble du domaine indo-européen) et *$\widehat{k}$ (palatale conservée dans l'aire *satəm* et confondue avec la vélaire pure dans les langues *centum*). Nous nous rallions à la première thèse (Lamberterie, 1980; Létoublon-Lamberterie, 1980, 317-8; Lamberterie, 1986, 53-7), estimant, avec Meillet, qu'il est simpliste de restituer de l'indo-européen en additionnant ce qu'on trouve dans les langues indo-européennes. On a cherché des traces des labiovélaires en albanais et en indo-iranien (Bolognesi, *l.c.*, avec bibl.); les exemples allégués ne sont guère convaincants, mais ce n'est pas ici le lieu de développer ce point.

c) La mutation consonantique : innovation ou archaïsme ?

En ce qui concerne le mode d'articulation des occlusives, les comparatistes ont longtemps reconstruit l'indo-européen en projetant dans la préhistoire le système indien, avec ses quatre modes d'articulation : sonores simples, sonores aspirées, sourdes simples et sourdes aspirées[42]. En réalité, la série des sourdes aspirées ne peut être mise sur le même plan que les autres, car elle n'apparaît que sur une aire dialectale limitée et dans des conditions particulières. On ne peut reconstruire de sourdes aspirées que sur la base des correspondances gr. χ = arm. *x* (ach-Laut) = skr. *kh*, gr. ϕ = arm. *p'* = skr. *ph*, gr. θ = arm. *t'* = skr. *th*, lesquelles ne se rencontrent guère que dans des termes expressifs et/ou sous l'influence d'un **s* précédant l'occlusive; plutôt que des phonèmes, les sourdes aspirées sont des variantes des sourdes simples, comme le montre l'alternance **k̑/kh* (> arm. *s/x*), **k/kh* (> arm. *k'/x*), **t/th*, **p/ph* (> arm. *h/p'*), fréquente dans le lexique des trois langues en question[43]. Cette isoglosse entre le grec, l'arménien et l'indo-iranien s'ajoute à celles qui ont été signalées plus haut. Si l'on remonte à l'indo-européen commun, seules les trois premières séries sont à considérer.

Selon la théorie classique, telle qu'elle est développée notamment dans l'*Esquisse* de Meillet, l'arménien et le germanique ont fait subir au système des occlusives indo-européennes un bouleversement complet, connu sous le nom de *mutation consonantique*[44] : les sonores aspirées indo-européennes sont devenues des sonores simples, les sonores simples des sourdes simples et les sourdes simples des sourdes aspirées, susceptibles d'aboutir ultérieurement à des spirantes. Il s'agit là d'innovations indépendantes, qui n'impliquent aucune affinité particulière entre le germanique et l'arménien.

Ces vues traditionnelles ont été contestées sur deux points :

1) Un premier point a trait à la série des « sonores aspirées » indo-européennes. D'après É. Benveniste et H. Vogt, les consonnes *b*, *d*, *g*, *j*, *ǰ* de l'arménien classique auraient été non pas des sonores simples, mais des aspirées, conservées sans changement depuis l'indo-européen (exception faite de la palatalisation pour *j* et *ǰ*) et comparables à la série *bh*, *dh*, *h* (< indo-iranien **jh*), *gh* du sanskrit[45]. Cette analyse, si l'on pouvait

42. Voir en ce sens notamment Brugmann-Delbrück, 1897-1916, I/2 (1897), 623-721.

43. Voir Meillet, 1922, 78-83; 1977, 274-85 (= *BSL* 36, 1935, 109-20); Hiersche, 1964; Lamberterie, 1982; 1990, 823-4 et 841-2.

44. Meillet, 1936, 25-38. Cf. aussi 1922, 89-96 (reproduit dans 1977, 391-97).

45. Benveniste, 1959, 46-56, avec histoire de la question (l'idée remonte à Pedersen, 1982, 114-5 = *ZVS* 39, 1906 [écrit en 1904], 336-7) et 1961; Vogt, 1988, 430-48 (= *NTS* 18, 1958, 143-61) et 458-61 (= *Vopr. Jaz.* 1961, 3, 40-3). Cf. aussi, dans le même sens, Lamberterie, 1973-74.

l'accepter, serait de grande portée, car elle sortirait de leur isolement les sonores aspirées du sanskrit et prouverait que la mutation consonantique de l'arménien n'a altéré que partiellement le système des occlusives indo-européennes ; mais les données de l'arménien moderne sur lesquelles elle s'appuie peuvent s'interpréter autrement[46].

2) Bon nombre de linguistes ont contesté, au nom de la vraisemblance typologique, le système reconstruit pour l'indo-européen, car il ne semble correspondre à aucun modèle connu[47]. Une série d'études convergentes, dont l'initiateur a été A. Martinet, tend à considérer que les sonores non aspirées [**b*], **d*, **g*, **g*^{w} que l'on reconstruit en indo-européen étaient en réalité des glottalisées ; dans cette perspective, les éjectives *p'*, *t'*, *k'*, *c'*, *č'* attestées dans une partie des dialectes arméniens modernes orientaux ne résulteraient pas d'une influence des langues du Caucase, mais continueraient directement une série indo-européenne[48]. Allant plus loin, certains comparatistes – ainsi P. Hopper, T. Gamkrelidze et V.V. Ivanov – ont proposé de substituer au schéma traditionnel indo-européen **d* (sonores) – **t* (sourdes) – **dh* (sonores aspirées) un système **t'* (glottalisées) – **t/th* (sourdes) – **d/dh* (sonores), en admettant que les sonores et les sourdes pouvaient avoir deux réalisations, aspirée et non aspirée[49].

Il s'agit là d'une reconstruction entièrement nouvelle, qui prend exactement le contrepied des conceptions traditionnelles[50]. Dans la théorie classique, en effet, on admet que la majorité des langues de la famille ont conservé, pour l'essentiel, le système ancien (parfois en le simplifiant), à l'exception du germanique et de l'arménien, qui sont passés par une mutation consonantique. Dans la nouvelle théorie, ce sont au contraire les langues à mutation consonantique que l'on tient pour conservatrices, ou, plus exactement, la prétendue mutation se révèle n'être qu'un mirage : on considère que le germanique et l'arménien ont gardé l'articulation sourde des glottalisées, ce qui, par contrecoup, a imposé l'allophone aspiré dans

---

46. Voir Pisowicz, 1976, 25 sq. ; Kortlandt, 1978a ; Lamberterie, 1984, 226-30.

47. Voir notamment Jakobson, 1962, 523-32 (première publication en 1958).

48. Voir Martinet, 1953, 70 n. 1 ; Gamkrelidze-Ivanov, 1973 ; 1984, 41-54 ; Haudricourt, 1975, 267-72 ; Hagège-Haudricourt, 1978, 123-5 et 165-6 ; Gamkrelidze, 1984 et 1985. F. Kortlandt (1978a, 12-5 ; 1978b, 107 et 110-1) estime que les éjectives de l'arménien sont bien héritées, mais reconstruit en indo-européen des glottalisées sonores.

49. Voir notamment Gamkrelidze-Ivanov, 1973 ; 1984, 5-80 ; Hopper, 1973, 1981, 1982 ; Haudricourt, 1975 ; Bomhard, 1979 ; Kortlandt, 1978b et 1985 (cette dernière étude faisant des concessions à la théorie classique) ; Gamkrelidze, 1981, 1984, 1987, 1989, 1990. Synthèse utile chez Bomhard, 1988, avec une riche bibliographie. État de la question et discussions récentes chez Vennemann (ed.), 1989, 83-263.

50. T. Gamkrelidze parle, en toute modestie, de « révolution copernicienne » (1987 et 1990).

les sourdes quand le trait de glottalisation s'est trouvé menacé dans la série précédente, et, par un second contrecoup, l'allophone non aspiré dans les sonores [51]. S'il faut parler de mutation consonantique, c'est bien plutôt à propos des autres langues (indo-iranien, etc.), car elles ont changé en sonores les anciennes glottalisées sourdes, ce qui, par contrecoup, a entraîné la réalisation aspirée de la série sonore. L'arménien apparaît alors comme une langue particulièrement archaïque, trait lié à la théorie de T. Gamkrelidze et V.V. Ivanov selon laquelle le berceau des Indo-Européens se serait situé précisément en Arménie.

Sans entreprendre ici une discussion d'ensemble de cette théorie, nous nous limiterons à quelques remarques concernant les données arméniennes [52] :

a) Le point le plus solide de la théorie est la réinterprétation des sonores simples de l'indo-européen comme d'anciennes glottalisées, mais il ne s'ensuit pas que les éjectives de certains dialectes arméniens modernes orientaux continuent *directement* une série indo-européenne : quand on connaît l'importance des phénomènes de glottalisation sur l'aire caucasienne, une telle idée paraît contestable, et le traitement **dw- > erk-* (ci-dessus 3a) invite à penser que le **t'* du plus ancien indo-européen était passé à **d* dans la préhistoire de l'arménien comme dans le reste du domaine; rien ne prouve d'ailleurs qu'il faille reconstruire les glottalisées indo-européennes comme des sourdes et non comme des sonores.

b) Le traitement i.-e. **k̂* > arm. *s* (*sirt* « cœur » < **k̂ērdi*) prouve qu'on ne peut poser en indo-européen une série **k/kh* de sourdes potentiellement aspirées dont l'arménien attesterait la réalisation aspirée. Tout s'explique, au contraire, si l'on admet que dans la préhistoire de la langue il s'est produit d'abord une palatalisation **k* > *k̂*, commune à l'aire *satəm*, puis, ensuite seulement, une mutation qui n'a pu atteindre un **k̂* > **ś* déjà sorti du système des occlusives : la mutation n'a porté que sur un résidu et non sur l'ensemble du système indo-européen. Il y a là une isoglosse notable avec l'indo-iranien, puisque **k̂* aboutit en sanskrit à la sifflante palatale *ś*, donc à une non occlusive, au lieu que **ĝ* et **ĝh* aboutissent aux mi-occlusives *j* et **jh* > *h*, c'est-à-dire au même résultat que dans la préhistoire de l'arménien avant la mutation qui a fait passer **j* à *c* (= /*ts*/) et **jh* à *j* (= /*dz*/). L'évolution ultérieure de **ś* en *s* est attestée non seulement en

51. Pour le germanique, voir en ce sens Normier, 1977 et Vennemann, 1984.

52. Voir Pisowicz, 1988; état de la question un peu plus ancien chez Lamberterie, 1984, 226-30. Sur l'ensemble de la théorie, objections de Dunkel, 1981 (d'où une vive polémique avec Hopper, dont on trouve un écho chez Hopper, 1981, 140-1); Haider, 1985; Huld, 1986; Mayrhofer, 1986, 96-7; Szemerényi, 1987, I, 397-410 (= *TPS* 1985, 3-15); Djahukian, 1990.

arménien, mais aussi dans la plupart des parlers iraniens, sans parler de sl. *s* en regard de balt. *š*.

c) C'est un paradoxe que de tenir pour conservatrices des langues comme le germanique ou l'arménien, car elles manifestent encore au cours de leur histoire une tendance persistante au chamboulement des séries : le nom *Tigran* de l'arménien ancien, conservé dans les dialectes orientaux, aboutit en arménien moderne occidental à *Dikran*, par l'effet d'une mutation comparable à celle qui sépare le haut-allemand des autres dialectes germaniques. Dans le cadre de la théorie classique, cette tendance s'interprète, tout naturellement, comme la continuation d'un processus entamé dans la préhistoire de la langue. Les autres langues indo-européennes, au contraire, font preuve sur ce point d'une bien plus grande stabilité au cours de leur histoire; il est donc raisonnable de les tenir pour conservatrices en ce qui concerne également leur préhistoire. Comme le système consonantique de l'arménien classique ressemble d'une manière frappante à celui des langues non indo-européennes du Caucase du sud (géorgien, et plus encore mingrélien), l'idée, chère à Meillet, selon laquelle les mutations du germanique et de l'arménien seraient dues à l'influence d'un substrat reste plausible aujourd'hui encore.

d) La prothèse vocalique

Il s'agit là d'une isoglosse qui unit l'arménien au grec et aussi au phrygien, comme on le voit par le nom de l'« homme » : arm. *ayr* (nom.) < **anēr* = gr. ἀνήρ (gén. *aṙn* < **anros* = ἀνδρός), phryg. αναρ, en regard de skr. *nar-/nr̥-*. Dans cet exemple, la prothèse provient du dégagement d'une voyelle consécutif à la disparition d'une consonne « laryngale » **h_2*, consonne qui remonte à l'état le plus ancien de l'indo-européen et dont une autre trace indirecte est l'allongement du composé skr. *sū-nára-* « à la force puissante » < **$h_1$su-$h_2$nór^e/$_o$-*. Mais les voyelles prothétiques ne remontent pas toutes à des laryngales; certaines d'entre elles ont une origine récente ou tiennent à des contraintes du système phonétique de la langue. L'arménien partage notamment avec le grec la particularité de ne pas admettre la consonne **r-* en début de mot, sans qu'on puisse savoir si c'est là un archaïsme ou une innovation [53].

---

53. Discussion chez Lamberterie, 1984, 230-1 (à propos des thèses d'É. Benveniste). La prothèse vocalique de l'arménien et du grec a été l'objet d'une multitude de travaux : voir Olsen, 1985, avec histoire de la question.

4. Morphologie

On a déjà signalé plus haut (§ 2) que le grec, le phrygien, l'arménien et l'indo-iranien constituaient un ensemble que l'on peut définir comme l'*aire de l'augment*, aire qui coïncide avec celle de l'aoriste thématique. Encore faut-il préciser qu'à la date tardive où l'arménien nous est connu, la distinction entre thèmes thématiques et thèmes athématiques n'est plus pertinente dans la flexion de l'aoriste : tant à l'actif qu'au moyen, la langue n'a qu'un seul jeu de désinences verbales. Quant à l'augment, ce n'est ni une marque facultative comme en grec homérique et en sanskrit védique, ni une marque obligatoire comme en sanskrit et en grec classiques, mais un élément à fonction purement rythmique dont l'emploi est réglé par la structure de la racine; l'augment n'apparaît qu'à certaines formes de l'aoriste pour donner plus de consistance à des formes qui sans cette adjonction seraient monosyllabiques : ainsi, du thème d'aoriste *ber-* «porter», *e-ber* «il porta» (avec désinence -ϕ) en regard de *ber-i* «je portai» ou de *ber-aw* «il fut porté». C'est là le résultat d'une longue évolution dont le détail nous échappe mais qui n'est pas sans évoquer, à certains égards, l'emploi de l'augment en grec moderne (έδεσα «je liai» *vs* δέθηκα «je fus lié»). D'une manière générale, la structure du verbe arménien classique, qui est fondée sur l'articulation de deux thèmes verbaux, le présent et l'aoriste, et dont le pivot est soit l'aoriste (dans les formations héritées), soit le présent (dans les formations récentes), rappelle fortement celle du grec moderne. Il s'agit là de développements parallèles dans deux langues proches l'une de l'autre.

Si le système verbal résulte d'une série d'innovations qui sont allées dans le sens d'une simplification, l'arménien a, en revanche, gardé une riche flexion nominale, avec sept cas, ce qui est surprenant pour une langue qui dans sa préhistoire a perdu les finales indo-européennes. Ce maintien d'une structure archaïque est peut-être dû à l'influence des langues environnantes, notamment celles du Caucase du sud, qui ont une riche déclinaison; ce qui invite à chercher en ce sens, c'est que l'arménien ignore tout autant que ces langues le genre grammatical, trait qui, pour une langue indo-européenne, représente évidemment une innovation.

5. Lexique

Les travaux classiques de Pedersen et de Meillet (ci-dessus § 2) restent, pour l'essentiel, valables aujourd'hui encore : c'est avec le grec que le

vocabulaire arménien offre les correspondances les plus nombreuses et les plus frappantes. Parmi les éléments du lexique limités à ces deux langues, on peut citer, par exemple, les deux racines homonymiques arm. *awel-* = gr. ὀφελ- «accroître» et arm. *awel-* = gr. ὀφελ- «balayer», qui n'en faisaient sans doute qu'une seule à l'origine [54].

Parfois la prise en compte de l'arménien permet d'éclairer l'origine des données grecques. Ainsi *ayr* «caverne» (homonyme de *ayr* «homme» = ἀνήρ), qui fait penser à gr. ἄντρον, ne peut y correspondre exactement; mais si l'on reconstruit en grec commun un substantif animé *ἀντήρ dont le collectif ἄντρα aurait entraîné secondairement la création d'un singulatif ἄντρον, on obtient une équation parfaite entre les deux langues et l'on restitue en grec une chaîne de formation identique à celle de ἀστήρ «astre» → ἄστρα → ἄστρον [55].

Dans le domaine de la formation des noms, où le grec présente tant de traits communs avec l'indo-iranien [56], on relève également nombre d'isoglosses gréco-arméniennes. L'une des plus remarquables est celle des deux adjectifs «long» de l'arménien, dont l'étymologie est due à Meillet [57] : d'une part arm. *erkar* = gr. δϝηρός < **dwā-ró-*, en regard du dérivé **dū-ró-* des langues périphériques (skr. *dūrá-*, lat. *dūrāre*), qui a le même suffixe mais une base radicale différente; et de l'autre arm. *erkayn* < **dwā-n-yo-*, dérivé (de date arménienne) en **-yo-* d'un thème **dwā-n-* reflété par l'adverbe gr. δϝήν «longtemps» (pour la formation, cf. arm. *layn* «large» < **plā-nyo-* en regard de lat. *plānus*); dans les deux langues, il faut partir de l'éclatement d'un thème hétéroclitique en **-r-/-n-*.

Il arrive que l'on puisse restituer des mots gréco-arméniens d'une certaine ampleur. Ainsi gr. δρεπάνη «faucille» a un répondant exact dans arm. *artewan* < **drepanā* «sourcil» (désignation métaphorique) [58]. Comme il existe en grec un certain nombre de noms d'instruments en -ανον et en -άνη, on serait tenté *a priori* de considérer δρεπάνη comme un dérivé de date grecque formé sur le verbe δρέπω «cueillir», et du côté arménien le suffixe *-an-* est productif également pour former des noms d'instruments [59];

---

54. Voir en ce sens Pedersen, 1982, 114 (= *ZVS* 39, 1906, 336). Dans une étude récente (Lamberterie, 1992), nous montrons que la comparaison avec l'arménien permet de clarifier le dossier complexe des trois verbes gr. ὀφέλλω («balayer – accroître – devoir»), auquel le déchiffrement du linéaire B a donné une orientation nouvelle.

55. Voir Lamberterie, 1978, 243-5.

56. Voir sur ce point Euler, 1979.

57. Meillet, 1977, 209-12 (= *REArm* 4, 1924, 1-4), avec nos prolongements 1988, 224-5.

58. Voir Lamberterie, 1983.

59. Cf. le substantif *gawazan* «houlette» cité ci-dessus § 1b. Sur le suffixe arménien, voir Bolognesi, 1990a, 225-38 (= *RL* 5, 1962, 105 sq.); sur les données grecques, Chantraine, 1933, 196-200.

mais il est exclu qu'il s'agisse, en l'occurrence, de créations parallèles. C'est dire que ces suffixes ont derrière eux une longue histoire.

Bibliographie

ADJARIAN (*Ačaṙyan*) H. (1909). – *Classification des dialectes arméniens*, Paris.

– (1971 -1979). – *Hayeren armatakan baṙaran* [Dictionnaire étymologique de l'arménien], Erevan. I (*A-D*), 1971 ; II (*E-K*), 1973 ; III (*H-Č'*), 1977 ; IV (*P-F*), 1979 (édition imprimée de l'original paru en polycopie en 7 vol., de 1926 à 1935).

AWETIK'EAN G., SIWRMELEAN X., AWGEREAN M. (1836-1837). – *Nor baṙgirk' haykazean lezui* [Nouveau Dictionnaire de la langue arménienne], Venise. I (*A-K*), 1836 ; II (*H-F*), 1837.

BADER F. (1973). – « Lat. *nempe, porceo* et les fonctions des particules pronominales », *BSL* 68, 27-75.

– (1985). – « De skr. *anyá-* à *ā́rya-* : noms i.-e. de l'"autre" », *BSL* 80, 57-90.

BENVENISTE É. (1930). – « Sur *-r-* de arménien *merk* "nu" », *REArm* 10, 187.

– (1959). – « Sur la phonétique et la syntaxe de l'arménien classique », *BSL* 54, 46-68.

– (1961). – Проблемы армянского консонантизма, ВЯ (*Vopr. Jaz.*) 3, 37-39.

BOLOGNESI G. (1960). – *Le fonti dialettali degli imprestiti iranici in armeno*, Milan.

– (1984). – « Arménien et iranien dans l'œuvre d'É. Benveniste », *É. Benveniste aujourd'hui, Actes du Colloque international du CNRS (Tours, 28-30 septembre 1983)*, II, 183-196.

– (1987). – « Il contributo di Antoine Meillet agli studi di linguistica armena », *L'opera scientifica di Antoine Meillet*, ed. A.Q. Moreschini, Pise, 119-146.

– (1988). – « La stratificazione del lessico armeno nell'opera di J.J. Schröder », *Energeia und Ergon. Festschrift für E. Coseriu*, edd. J. Albrecht, J. Ludtke et H. Thun, Tübingen, II, 561-571.

– (1990a). – *Studi glottologici filologici orientali*, Brescia.

– (1990b). – « Arménien et iranien. Une théorie linguistique passée inaperçue », *When Worlds Collide : The Indo-Europeans and the Pre-Indo-Europeans*, edd. T.L. Markey et J.A.C. Greppin, Ann Arbor, 227-236.

BOMHARD A.R. (1979). – « The Indo-European phonological system : new thoughts about its reconstruction and development », *Orbis* 28, 66-110.

– (1988). – « Recent trends in the reconstruction of the Proto-Indo-European consonant system », *HS* 101, 2-25.

BONFANTE G. (1937). – « Les isoglosses gréco-arméniennes. I. Faits phonétiques », *Mélanges... Pedersen*, Aarhus, 125-33.

– (1946). – « Armenians and Phrygians », *Armenian Quarterly* I/1, 82-97.

– (1978). – « Faits phonétiques romans éclairant des faits indo-européens », *BSL* 73, 221.

– (1982). – « The place of Armenian among the Indo-European languages », *AION* 14, 151-169.

BOPP F. (1816). – *Über das Conjugationssystem der Sanskritsprache in Vergleichung mit jenem der griechischen, lateinischen, persischen und germanischen Sprache*, Francfort-sur-le-Main (repr. Hildesheim – New York, 1975).

– (1833-1849). – *Vergleichende Grammatik des Sanskrit, Zend, Griechischen, Lateinischen, Litthauischen, Gothischen und Deutschen*, Berlin. I, 1833 ; II, 1835 (avec adjonction de *Altslawischen* dans le titre) ; III, 1837 ; IV, 1842 ; V, 1849.

– (1857-1861). – *Vergleichende Grammatik des Sanskrit, Ṣend, Armenischen, Griechischen, Lateinischen, Litauischen, Altslawischen, Ġotischen und Deutschen*, Berlin. I, 1857 ; II, 1858 ; III, 1861 (+ index par C. Arendt, 1863).

– (1866-1872). – *Grammaire comparée des langues indo-européennes, comprenant le sanscrit, le zend, l'arménien, le grec, le latin, le lithuanien, l'ancien slave, le gothique et l'allemand*, traduite sur la deuxième édition par M. Bréal, Paris. I, 1866 ; II, 1868 ; III, 1869 ; IV, 1872 (+ V, index par F. Meunier, 1874).

BRIXHE C. (1983). – « Épigraphie et grammaire du phrygien : état présent et perspectives », *Le lingue indoeuropee di frammentaria attestazione (Atti del Convegno della Società Italiana di Glottologia e della Indogermanische Gesellschaft, Udine*, 22-24 set. 1981), Pise, 109-133.

BRIXHE C. – LEJEUNE M. (1984). – *Corpus des inscriptions paléo-phrygiennes*, Paris.

BRUGMANN K. – DELBRÜCK B. (1897-1916). – *Grundriß der vergleichenden Grammatik der indogermanischen Sprachen*, Strasbourg. I-II, 2e éd. (K.B.), 1897-1916 ; III-V (B.D.), 1893-1900.

CHANTRAINE P. (1933). – *La Formation des noms en grec ancien*, Paris.

DJAHUKIAN. – *voir* JAHUKYAN.

DUMÉZIL G. (1948). – « L'Arménien à l'École Nationale des Langues Orientales Vivantes », *Cent-cinquentenaire de l'École des Langues Orientales*, Paris, 65-74.

DUNKEL G. (1981). – « Typology *versus* reconstruction », *Bono Homini Donum : Essays in Historical Linguistics, in Memory of J. Alexander Kerns*, edd. Y.L. Arbeitman et A.R. Bomhard, Amsterdam, II, 559-569.

EULER W. (1979). – *Indoiranisch-griechische Gemeinsamkeiten der Nominalbildung und deren indogermanische Grundlage*, Innsbruck.

GAMKRELIDZE T.V. (1981). – « Language typology and language universals and their implication for the reconstruction of the Indo-European stop system », *Bono Homini Donum : Essays ... Kerns* (cf. Dunkel, 1981), II, 571-609.

– (1984). – « Индоевропейская "глоттальная теория" и система древнеармянского консонантизма", *Miǰazgayin hayerenagitakan gitažołov (Erevan, 21-25 sept. 1982), Zekuc'umner*, Erevan, 31-34.

– (1985). – « The Indo-European "glottalic theory" and the system of Old Armenian consonantism », *Studia Linguistica Diachronica et Synchronica, Werner Winter sexagenario anno MCMLXXXIII*, edd. U. Pieper et G. Stickel, Berlin – New York – Amsterdam, 281-284 (= 1984).

– (1987). – « The Indo-European glottalic theory : a new paradigm in i.-e. comparative linguistics », *JIES* 15, 48-59.

– (1989). – « Markedness, sound change and linguistic reconstruction », *Markedness in Synchrony and Diachrony*, ed. O.M. Tomić, Berlin – New York, 1989, 87-101.

– (1990). – « The Indo-European glottalic theory in the light of recent critique », *FoLH* 9/1, 3-12.

GAMKRELIDZE T.V. – IVANOV V.V. (1973). – « Sprachtypologie und die Rekonstruktion der gemeinindogermanischen Verschlüsse », *Phonetica* 27, 150-156.

– (1984). – Индоевропейский язык и Индоевропейцы, 2 vol., Tbilissi.

GODEL R. (1975). – *An Introduction to the Study of Classical Armenian*, Wiesbaden.

– (1982). – *Linguistique arménienne : Études diachroniques*, Vaduz-Paris.

GREPPIN J. – KHACHATURIAN A. (1986). – *A Handbook of Armenian Dialectology*, Delmar (NY).

HAAS O. (1966). – *Die phrygischen Sprachdenkmäler*, Sofia.

HAGÈGE Cl. – HAUDRICOURT A.G. (1978). – *La Phonologie panchronique*, Paris.

HAIDER H. (1985). – « The fallacy of typology », *Lingua* 65, 1-27.

HAMP E.P. (1979). – « Indo-European *$*g^{w}en\text{-}H_a$* », *ZVS* 93, 1-7.

– (1983). – « Philologica varia », *REArm* 17, 5-12.

HAUDRICOURT A.G. (1975). – « Les mutations consonantiques (occlusives) en indo-européen », *Mélanges linguistiques offerts à Émile Benveniste*, Paris-Louvain, 267-272.

HIERSCHE R. (1964). – *Untersuchungen zur Frage der Tenues Aspiratae im Indogermanischen*, Wiesbaden.

HOPPER P.J. (1973). – « Glottalized and murmured occlusives in Indo-European », *Glossa* 7, 141-166.

– (1981). – « "Decem" and "Taihun" languages : an Indo-European isogloss », *Bono Homini Donum : Essays ... Kerns* (cf. Dunkel, 1981), I, 133-142.

– (1982). – « Areal typology and the early Indo-European consonant system », *The Indo-Europeans in the Fourth and Third Millenia*, ed. E.C. Polomé, Ann Arbor, 121-139.

HÜBSCHMANN H. (1897). – *Armenische Grammatik*. Erster Teil : *Armenische Etymologie*, Leipzig.

– (1976). – *Kleine Schriften zum Armenischen*, ed. R. Schmitt, Hildesheim – New York.

HULD M. (1986). – « On the unacceptability of the Indo-European voiced stops as ejectives », *IF* 91, 67-78.

JAHUKYAN (Джаукян) G.B. (1978). – **Общее и армянское языкознание**, Erevan.

– (1980). – « On the position of Armenian in the Indo-European languages », *First International Conference on Armenian Linguistics, Proceedings*, ed. J. Greppin, Delmar (NY), 3-16.

– (1982). – **Сравнительная грамматика армянского языка**, Erevan.

– (1983). – « On etymological doublets and parallels of Armenian », *AArmL* 4, 23-34.

– (1984). – « **Об этимологических дублетах и параллелях в армянском языке**, *Miǰazgayin hayerenagitakan gitažołov* (*Erevan, 21-25 sept. 1982*), *Zekuc'umner*, Erevan, 146-160 (= 1983).

– (1985). – « Die archaischen Schichten armenischer Dialektwörter », *Studia Linguistica Diachronica et Synchronica, Werner Winter sexagenario anno MCMLXXXIII*, edd. U. Pieper et G. Stickel, Berlin – New York – Amsterdam, 151-160.

– (1987). – *Hayoc' lezvi patmut'yun. Naxagrayin žamanakašrǰan* [Histoire de la langue arménienne : époque prélittéraire], Erevan.

– (1990). – « A variational model of the Indo-European consonant system », *HS* 103, 1-16.

JAKOBSON R. (1962). – *Selected Writings*, I, La Haye.

JANKOWSKY K.R. (1972). – *The Neogrammarians*, La Haye – Paris.

KLINGENSCHMITT G. (1972). – *Das altarmenische Verbum*, Wiesbaden.

KORTLANDT F. (1978a). – « Notes on Armenian historical phonology II (The second consonant shift) », *Stud. Caucas.* 4, 9-16.

– (1978b). – « Proto-Indo-European obstruents », *IF* 83, 107-118.

– (1985). – « Proto-Indo-European glottalic stops : the comparative evidence », *FLH* VI/2, 183-201.

– (1986). – « Armenian and Albanian », *La Place de l'arménien dans les langues indo-européennes*, edd. M. Leroy et F. Mawet, Louvain, 38-47.

– (1989). – « The making of a puzzle », *AArmL* 10, 43-52.

LAMBERTERIE Ch. de (1973-74). – « Les occlusives sonores aspirées de l'arménien classique », *REArm* 10, 39-44.

– (1978). – « Armeniaca I-VIII : études lexicales », *BSL* 73, 243-285.

– (1978-79). – « Armeniaca X : une isoglosse gréco-arménienne », *REArm* 13, 31-40.

– (1980). – « Échange de gutturales en arménien » (= Armeniaca XII), *AArmL* 1, 23-37.

– (1982). – « La racine **(s)per-ĝh-* en arménien » (= Armeniaca XVI), *REArm* 16, 57-68.

– (1983). – « Une bagatelle étymologique (*artewanunk'* / δρεπάνη) » (= Armeniaca XVII), *REArm* 17, 21-22.

– (1984). – « Benveniste et la linguistique arménienne », *É. Benveniste aujourd'hui, Actes du Colloque international du CNRS (Tours, 28-30 sept. 1983)*, Paris, II, 225-238.

– (1985). – « Arménien *apak'ēn* » (= Armeniaca XVIII), *AArmL* 6, 39-44.

– (1986). – « Deux isoglosses entre l'arménien et l'indo-iranien » (= Armeniaca XIX-XX), *La Place de l'arménien* (cf. Kortlandt, 1986), 48-61.

– (1988). – « Meillet et l'arménien », *HEL* 10/2 (*Antoine Meillet et la linguistique de son temps*), 217-234.

– (1989). – « Arménien *ari* et *anari* », *Études irano-aryennes offertes à Gilbert Lazard*, edd. C.-H. de Fouchécour et Ph. Gignoux, Paris, 237-246.

– (1990). – *Les Adjectifs grecs en -υς*, 2 vol., Louvain-la-Neuve.

– (1992). – « Le problème de l'homonymie : les trois verbes ὀφέλλω en grec ancien », *La langue et les textes en grec ancien* (Actes du colloque Pierre Chantraine, Grenoble, 5-8 septembre 1989), ed. F Létoublon, Amsterdam, 201-217.

LÉTOUBLON F. – LAMBERTERIE Ch. de (1980). – « La roue tourne », *RPh* 54, 305-326 (pp. 314-326 = C.L. Armeniaca XIII).

MARTINET A. (1953). – « Remarques sur le consonantisme sémitique », *BSL* 49, 67-78.

MAYRHOFER M. (1986). – *Indogermanische Grammatik.* I/2 : *Lautlehre*, Heidelberg.

MEILLET A. (1913). – *Altarmenisches Elementarbuch*, Heidelberg.

– (1922). – *Les Dialectes indo-européens*, 2e éd. (nouveau tirage avec une introduction nouvelle), Paris (1re éd., 1908).

– (1925). – *La Méthode comparative en linguistique historique*, Oslo.

– (1936). – *Esquisse d'une grammaire comparée de l'arménien classique*, 2e éd., Vienne (1re éd., 1903).

– (1937). – *Introduction à l'étude comparative des langues indo-européennes*, 8e éd., Paris (repr. Alabama, 1964 – 1re éd., 1903).

– (1962). – *Études de linguistique et de philologie arméniennes*. I, Lisbonne.

– (1977). – *Études de linguistique et de philologie arméniennes*. II, Louvain.

NORMIER R. (1977). – « Idg. Konsonantismus, germ. "Lautverschiebung" und Vernersches Gesetz », *ZVS* 91, 171-218.

OLSEN B. (1985). – « On the development of Indo-European prothetic vowels in classical Armenian », *REArm* 19, 5-17.

– (1989). – « A trace of Indo-European accent in Armenian », *HS* 102, 220-240.

PEDERSEN H. (1925). – *Le Groupement des dialectes indo-européens*, Copenhague.

– (1982). – *Kleine Schriften zum Armenischen*, ed. R. Schmitt, Hildesheim – New York.

PETERMANN H. (1837). – *Grammatica linguæ armeniacæ*, Berlin.

PICARD M. (1989). – « A reanalysis of Armenian prothesis and metathesis », *FLH* 10/1-2, 61-69.

PISOWICZ A. (1976). – *Le développement du consonantisme arménien*, Cracovie.

– (1988). – « Objections d'un arménologue contre la théorie glottale », *FO* 25, 213-225.

PORZIG W. (1954). – *Die Gliederung des indogermanischen Sprachgebiets*, Heidelberg (repr., 1974).

SCHLEICHER A. (1861-1862). – *Compendium der vergleichenden Grammatik der indogermanischen Sprachen*, Weimar (3e éd., 1871).

SCHMIDT K.H. (1980). – « Armenian and Indo-European », *First International Conference on Armenian Linguistics : Proceedings*, ed. J. Greppin, Delmar (NY), 35-58.

– (1990). – « The Indo-European basis of Proto-Armenian : Principles of Reconstruction », *AArmL* 11, 33-47.

SCHMITT R. (1972) [74]. – « Die Erforschung des Klassisch-Armenischen seit Meillet (1936) », *Kratylos*, 17, 1-68.

– (1975). – « Von Bopp bis Hübschmann : Das Armenische als indogermanische Sprache », *ZVS* 89, 3-30.

– (1981). – *Grammatik des Klassisch-Armenischen mit sprachvergleichenden Erläuterungen*, Innsbruck.

– (1983). – « Iranisches Lehngut im Armenischen », *REArm* 17, 73-112.

– (1986). – « Ein vergessener Pionier der Armenisch-Forschung », *o-o-pe-ro-si, Festschrift für Ernst Risch zum 75. Geburtstag*, ed. A. Etter, Berlin – New York, 650-653.

– (1987). – « Iranian influences in Armenian. I. General », *Encyclopaedia Iranica*, ed. E. Yarshater, Londres – New York, II, 445-459.

SCHRÖDER J.J. (1711). – *Thesaurus linguæ armeniacæ*, Amsterdam.

SOLTA G.R. (1960). – *Die Stellung des Armenischen im Kreise der indogermanischen Sprachen*, Vienne.

SZEMERÉNYI O. (1987). – *Scripta Minora*, 3 vol. (I : *Indo-European*; II : *Latin*; III : *Greek*), Innsbruck.

– (1991). – *Scripta Minora*, IV (*I.-E. Languages Other than Latin and Greek*), Innsbruck.

VENNEMANN T. (1984). – « Hochgermanisch und Niedergermanisch. Die Verzweigungstheorie der germanisch-deutschen Lautverschiebung », *PBB* 106, 1-45.

– (1986). – « Syllable-based sound changes in early Armenian », *AArmL* 7, 27-43.

– (1989). – T.V. (ed.), *The New Sound of Indo-European. Essays in Phonological Reconstruction*, Berlin – New York.

VERNER K. (1877). – « Eine ausnahme der ersten lautverschiebung », *KZ* 23, 97-130 (écrit en 1875).

VOGT H. (1988). – *Linguistique caucasienne et arménienne*, Oslo.

WATKINS C. (1963). – « Preliminaries to a historical and comparative analysis of the syntax of the Old Irish verb », *Celtica* 6, 1-49.

– (1969). – *Geschichte der indogermanischen Verbalflexion* (= *Indogermanische Grammatik*, ed. J. Kuryłowicz, III/1), Heidelberg.

WEITENBERG J.J.S. (1985). – « Additional *-n* in Armenian », *AArmL* 6, 101-106.

– (1986). – « Additional *h-*, initial *y-* and Indo-European **y-* in Armenian », *La Place de l'arménien* (cf. Kortlandt, 1986), 90-101.

WINDISCHMANN F. (1846). – « Die Grundlage des Armenischen im arischen Sprachstamme », *Abhandlungen der königlich bayerischen Akademie der Wissenschaften, philos.-philolog. Classe*, IV/2, Munich, 1-49.

WINTER W. (1966). – « Traces of early dialectal diversity in Old Armenian », *Ancient Indo-European Dialects*, edd. H. Birnbaum et J. Puhvel, Berkeley – Los Angeles, 201-211.

ZELLER H. (1927). – « Armenisch », *Die Erforschung der indogermanischen Sprachen* (= *Geschichte der indogermanischen Sprachwissenschaft* II), IV/2, Berlin – Leipzig, 85-104.

CHAPITRE 8

LE PHRYGIEN

Claude BRIXHE

0. Selon Hérodote (VII 73), les Phrygiens, au dire des Macédoniens, se seraient appelés Βρύγες tant qu'ils cohabitaient avec eux et vivaient en Europe, puis ils auraient changé leur nom en Φρύγες lorsqu'ils passèrent en Asie. Or, l'on sait que l'indo-européen **bh* aboutit à *ph* en grec, mais à *b* en phrygien; Φρύγες doit donc être le nom des Phrygiens en grec et Βρίγες/Βρύγες pourrait avoir été celui qu'ils se donnaient eux-mêmes, mais aussi celui que leur donnaient les Thraces et au moins une des composantes de la population macédonienne (où même traitement de l'aspirée).

0.1. Hérodote n'est pas le seul à faire venir les Phrygiens de la péninsule Balkanique, cf. encore Strabon, VII 3.2 et XII 8.3, ou Pline, V. XLI 145.

Du reste, certains récits mythiques de Macédoine ou de Thrace Occidentale rappelaient leur séjour dans la région [1]. En outre, il semble bien qu'ils aient laissé, de la Pélagonie aux abords de l'Athos, quelques tribus, connues sous divers noms (Βρύγοι, Βρίγες, Βρύκες, etc) [2] susceptibles d'être ramenés au même étymon.

---

1. Cf. «Les jardins de Midas» près du Bermion, Hérodote, VIII 138.

2. Cf. Vl. Georgiev, *Introduzione alla storia delle lingue indeuropee*, Rome 1966, 168. De telles variations désinentielles sont banales en grec pour les noms de peuples

Les Phrygiens d'Asie Mineure ou les populations résiduelles des Balkans sont parfois donnés comme thraces (*e.g.* Strabon, VII 3.2). On verra plus loin la réponse qu'apporte la langue à cette question.

0.2. On a longtemps associé la migration phrygienne avec la fin de l'empire hittite, vers −1200. Aujourd'hui, à la lumière des données archéologiques, on aurait tendance à la faire descendre jusqu'au Xe et même au IXe siècle avant J.-C [3].

S'il en était ainsi, faudrait-il voir un autre peuple que les Phrygiens dans les Muskis qui apparaissent dans les sources orientales – surtout assyriennes – à partir du XIIe siècle avant J.-C. et dans le nom desquels on avait cru reconnaître l'appellation des Phrygiens (à partir d'une de leurs composantes ?) auprès des Orientaux ? [4]

Toujours est-il que les Phrygiens entrent dans la littérature grecque dès Homère (*Iliade* II 862, III 184, XXIV 545), où ils apparaissent comme voisins orientaux et alliés des Troyens.

Au IXe et surtout au VIIIe siècle avant J.-C., ils constituent un empire puissant. Les fouilles effectuées dans leur capitale, Gordion, révèlent une société qui, par sa culture (mode de vie, architecture), n'est pas sans analogies avec celle des poèmes homériques. Les Phrygiens interviennent en Cilicie et en Cappadoce, où ils se heurtent aux Assyriens. Ils entretiennent des relations cultuelles et matrimoniales avec le monde grec. En −709, selon la chronique assyrienne, Midas (*Mita*, roi des *Muskis*) se soumet à Sargon II. Puis l'arrivée des Cimmériens, au début du VIIe siècle, le contraint au suicide selon la légende, le Grand Tumulus de Gordion correspondant peut-être à son tombeau. Que devient alors son empire ? Il pourrait s'être morcelé en petites principautés ; mais les Phrygiens, comme le montre l'extension de leurs monuments inscrits, continuent à occuper un vaste espace et sans doute à avoir une forte influence culturelle sur l'Anatolie. A la fin du VIIe siècle avant J.-C., ils tombent sous le joug lydien, relayé par le joug perse, sous lequel ils restent jusqu'à l'arrivée d'Alexandre en −333. Les Galates, vers −275, les repoussent vers le Sud. La présence romaine, à partir de la fin du IIe siècle avant J.-C., ne modifiera pas l'ambiance culturelle et linguistique grecque dans laquelle ils évoluent depuis la conquête macédonienne [5].

étrangers, cf. dans les Balkans précisément Ἠδωνοί/Ἠδῶνες/Ἠδῶναι pour une tribu thrace, ou Δαρδάνιοι/Δάρδανοι/Δαρδανεῖς/Δαρδανιᾶται pour l'ethnie illyrienne (? cf. «Le thrace» § 1.2.2) bien connue.

3. Voir I. Venedikov, *Pulpudeva* 2 (Plovdiv 1976), Sofia 1978, 163 ; Neumann, 16-17.

4. Voir, avec bibliographie, Brixhe, 1986, § 3.3, et 1987/1, § 5.3.

5. Esquisse de cette histoire : dans *Der kleine Pauly* 3, col. 1289, *s.v.* Midas II et III (D.O. Edzard et H.v. Geisau), et 4, col. 822-824, *s.v.* Phryger (G. Neumann) ; et chez Brixhe, 1986, §§ 3.1 à 3.3, 1987/1, §§ 5.3-5.4.

Jadis langue d'un peuple puissant et hégémonique, le phrygien est donc devenu le parler d'une population colonisée [6]. Ce changement de statut va naturellement limiter ses domaines d'utilisation.

1. Les documents

La langue des Phrygiens est documentée par deux groupes d'inscriptions actuellement séparés par trois siècles; cet hiatus ne cesse de se réduire au fil des découvertes épigraphiques.

1.1. Les textes paléo-phrygiens[7]

Ils ont été rassemblés par Cl. Brixhe et M. Lejeune (1984). À ce recueil, il faut à présent ajouter :

a) un sceau dont l'origine exacte est inconnue et qui se trouve dans un musée américain [8];

b) un graffite sur récipient d'argent (VIe siècle avant J.-C.), trouvé dans un tumulus à l'Ouest d'Uşak [9];

c) plusieurs graffites sur objets découverts dans un tumulus de la plaine d'Elmalı (Nord de la Lycie; fin du VIIIe ou VIIe siècle avant J.-C. ?) [10];

d) une épitaphe sur stèle de la fin du IVe siècle avant J.-C., vue sur le territoire de Dokimeion par Th. Drew-Bear;

e) une quinzaine de graffites inédits de Gordion.

---

6. Avec, pendant longtemps encore, suffisamment de dynamisme pour absorber éventuellement des éléments alloglosses, cf. l'anthroponyme galate Βωδορις dans l'épitaphe bilingue néo-phrygienne n° 34.

7. Les inscriptions paléo-phrygiennes, données en translitération latine, sont citées d'après Brixhe-Lejeune; les néo-phrygiennes, en caractères grecs, le sont d'après Haas, Brixhe, 1978, Brixhe-Waelkens et Brixhe-Neumann. Dans les pages qui suivent, on reconnaîtra donc aisément à la graphie l'appartenance d'une forme au paléo- ou au néo-phrygien.

8. O. Masson, *Kadmos* 26 (1987), 109-112.

9. Publié par R. Gusmani, *Studi di storia e di filologia anatolica dedicati a G. Pugliese Carratelli*, Florence 1988, 67-73; Cl Brixhe en montre le caractère authentiquement phrygien dans *Incontri Linguistici* 13 (1989-1990), 61-67.

10. E. Varinlioğlu, *Kadmos* 31 (1992), 10-20 et pl. I-II. Nous sommes alors proche de ce que Zgusta (carte *in fine*) appelle le «südphrygisch-lykisches Übergangsgebiet».

1.1.1.

Sur pierre, il s'agit le plus souvent d'inscriptions (fréquemment rupestres) appartenant au domaine cultuel ; mais T-01-03 ont vraisemblablement un contenu historique. Sur vases, on a généralement affaire à des noms indiquant la propriété. Le document le plus long comporte environ 285 signes (l'épitaphe inédite de Dokimeion susmentionnée) ; vient ensuite B-01 (façade rupestre de Bithynie) avec à peu près 230 signes. Les graffites se réduisent parfois à quelques lettres.

Ces inscriptions s'étalent des dernières décennies du VIIIe siècle au IIIe siècle avant J.-C. Elles ont été trouvées sur une aire très vaste :

1. en Phrygie occidentale, dont la célèbre « Cité de Midas » (W et M chez Brixhe-Lejeune) ;

2. en Bithynie occidentale (B) ;

3. dans la zone phrygo-galate (G et C, Gordion et environs) ;

4. en Ptérie, sur le site de l'ancienne capitale hittite, Hattusas/Boğazköy, et aux alentours (P) ;

5. en Tyanide (Cappadoce, T) ;

6. au Nord de la Lycie.

5 était une zone louvitophone et le phrygien y est arrivé pour des raisons historiques ponctuelles. **1** à **4** correspondent assurément à des régions phrygophones, ce qui éventuellement a pu être le cas de **6** également, à moins qu'il ne s'agisse là d'objets simplement amenés par la pratique du don et du contre-don.

À l'exception de l'épitaphe de Dokimeion (alphabet grec classique), les textes sont écrits, avec parfois signes d'interponction, en un alphabet semblable aux abécédaires grecs archaïques, avec un fonds de 17 lettres, auxquelles s'ajoutent quelques symboles spéciaux, dont ʅ/ʃ pour *j*[11] et ↑/T pour *ts*[12], non connus partout.

1.2. Les documents néo-phrygiens

Après une éclipse d'environ trois siècles, le phrygien resurgit au Ier siècle de notre ère[13], sous la forme, le plus souvent, de brèves imprécations[14] accompagnant fréquemment une épitaphe grecque et écrites

11. Avatar, comme le *iota* grec serpentin, du *jod* sémitique

12. Sans doute T diacrité, qui a peut-être fourni à l'ionien son *sampi*.

13. Le dernier document semble assignable au IIIe siècle après J.-C.

14. Du type « celui qui endommagera ce tombeau (ιος νι σεμουν κνουμανε κακουν αδδακετ), que celui-là soit... ». Le texte le plus long (l'épitaphe unilingue n° 116 = Brixhe-Neumann) comporte environ 290 signes.

en *scriptio continua* avec l'alphabet grec de l'époque. Aux 104 textes recensés par Brixhe, 1990, 60, n. 1 [15], on ajoutera deux inédits découverts par Th. Drew-Bear.

Un certain nombre de correspondances phonétiques, morphologiques, lexicales et syntaxiques assurent que les deux groupes de documents (paléo- et néo-phrygiens) reflètent deux états d'une même langue, cf *ios ni* (P-04) et ιος νι (*passim*), thème *kako-* dans *kakoioi/kakuioi* (G-02c, P-04b) et κακουν ou variantes (*passim*), *avtun/avtay* (W-01b) et αυτος (n° 33, 36), etc.

Devenu langue d'un peuple asservi, le phrygien est désormais confiné au domaine sacré.

Les incriptions ont été trouvées dans un périmètre délimité par Konya, la pointe Nord du Tuz Gölü, Eskişehir, Kütahya, Dinar. La plus forte densité s'observe sur la bordure Ouest et au N./N.-O. de la pointe Nord du Tuz Gölü. On signalera qu'avec le petit groupe découvert aux alentours de Laodicée Katakékauméné nous sommes aux confins de la Lycaonie et de la Phrygie. L'aire d'expression phrygienne semble donc s'être considérablement réduite depuis le paléo-phrygien.

1.3.

On a trouvé en 1926, à Üyücek sur le Haut Rhyndakos, à l'extrémité Ouest de la Phrygie, une inscription (Vᵉ/IVᵉ siècle avant J.-C. : Friedrich, 1932, 140-141, n° XI), dont on a cru qu'elle illustrait le mysien (μιξολύδιος/μιξοφρύγιος) évoqué par Strabon, XII 8.3. Il faut probablement y joindre un texte inédit découvert en Bithynie et écrit en un alphabet à peu près identique (Brixhe, 1975, 74; 1983, 130; 1990, 60-61). En fait, dans les deux cas, on a probablement affaire au phrygien ou à l'un de ses dialectes (ainsi Brixhe, 1975, *l.c.* et Neumann, 4).

1.4.

Enfin, pour apprécier le phrygien, nous pouvons utiliser deux autres sources :

1. Un certain nombre de gloses [16], livrées surtout par Hésychius, mais aussi parfois par la littérature grecque; certaines sont confirmées par les textes, ainsi βεκος «pain», Hérodote II 2, et néo-phr. n° 33, 76, 86, etc.; ζεμελεν· ἀνδράποδον, Φρύγες (Hésychius), sans doute vocatif d'un nom signifiant peut-être tout simplement «l'homme» (radical **dh(e)ghem-*

---

15. Haas, n° 1-110 (dont quelques documents grecs); Brixhe, 1978, n° 111-114; Brixhe-Waelkens, n° 115; Brixhe-Neumann, n° 116.

16. Elles sont rassemblées par Friedrich, 1941, col. 870-871, Haas, 157-172, et Neroznak, 133-154.

« terre »), en face du datif pluriel néo-phr. ζεμελως ou variantes (n° 21, 40, 42...) ; γλουρεα· χρύσεα, Φρύγες <καὶ> γλουρος· χρυσός (Hésychius), à côté de γλουρεος dans l'inscription inédite de Dokimeion signalée *supra* § 1.1 (radical **ghlōr-*, cf. grec χλωρός Brixhe 1982, 243-244 ; 1990, 93) ; etc.

2. Les inscriptions grecques de Phrygie a) qui nous fournissent quelques termes d'origine phrygienne, cf. τὸ βέννος « association des fidèles d'un dieu »[17] ; b) qui, par leur phonétique, leur morphologie et leur syntaxe nous donnent de précieuses indications sur le substrat/adstrat phrygien[18].

2. La langue

Le phrygien est encore mal compris ; mais, si les textes de quelque étendue restent opaques, si l'on est incapable d'entrevoir les divisions dialectales qui ne devaient pas manquer de parcourir un aussi vaste territoire[19], les efforts combinés de la comparaison et de l'analyse interne ont permis, dans les dernières décennies, de dessiner les grandes lignes de sa grammaire, sans, souvent, qu'on puisse préciser le sens de l'unité lexicale concernée.

2.1. La phonétique

2.1.1.

Le vocalisme[20] paléo-phrygien est relativement conservateur. Il a gardé les oppositions de quantité. Les voyelles longues héritées semblent y avoir conservé leur timbre, à l'exception de **ē*, qui s'est confondu avec

17. D'où βεννάρχης, βεννεύω, Ζεύς Βέννιος, voir Th. Drew-Bear, Chr. Naour, ANRW II 18.3 (1990), 1952-1991 (conclusions 1990-1991).

18. Voir Brixhe, 1983, 129, et surtout 1987, 110-116 et 157-158.

19. Cette incapacité pourrait être définitive, notamment pour le paléo-phrygien, si l'on continue à ne découvrir que des documents reflétant le plus haut niveau de la langue, comme ce fut le cas jusqu'ici hormis pour les graffites.

20. En dernier lieu, Brixhe, 1990.

*ā cf. le nominatif *matar* (W-04, etc.) en face de l'accusatif *materan* (W-01 a, etc.). Cette mutation a pu intervenir à la suite de l'insertion entre *ī et *ē d'un *ē* secondaire issu de la réduction de *ei* et de l'allongement compensatoire (plus tard, de la contraction). À l'arrière, le *ō* secondaire né dans les mêmes conditions risque d'avoir fusionné avec *ū, tandis que l'ancien *ō restait intact, cf. impératif *si↑eto* (W-08-10, avec finale *-to* = - τω du grec). Les voyelles brèves sont apparemment stables, sauf *o qui se ferme en *u* à la finale devant *n*, ainsi *avtun* (W-01b). La langue possède une double série de diphtongues en *-i* et *-u*, à premier élément bref ou long, cf. datifs *avtay* ou *materey* (*ibid*).

Les diphtongues à premier élément long perdent leur seconde composante peut-être dès la fin de la phase paléo-phrygienne, cf. le datif [*s*]*a soroi* (démonstratif + emprunt au grec σορός, épitaphe inédite de Dokimeion). Toujours est-il que c'est un fait généralisé en néo-phrygien, même si l'orthographe conserve encore parfois le souvenir de la situation ancienne, cf. σα σορου (n° 21) ou σα μανκα (nom du tombeau ou d'une de ses parties, n° 82), en face de σα ... σκελεδριαι (*idem*, n° 87) ou σαι... μανκαι (n° 35). Le phrygien a alors perdu les oppositions de quantité et a vraisemblablement ramené son système à cinq voyelles isochrones, *i, e, a, o, u*, auxquelles s'ajoutent quelques diphtongues, *au, ai, oi* (< **au*, **ai*, **oi*). Les longues moyennes se sont fermées, cf. κνουμανι (datif du nom du tombeau, avec finale issue de *-*ey*, *passim*), γλουρεα/γλουρος (*supra*) ou impératif 3[e] pers. sing. (quelle que soit l'étymologie du radical) ειτου (où -του < *-*tōd*, *passim*). En finale, *o* et *e* se ferment en *u* et *i*, cf. adverbe κακιν (n° 14) en face de κακε(ν) (n° 21, 88, etc.).

2.1.2.

On a longtemps prétendu [21] que le système consonantique phrygien était dominé par la mutation (*Lautverschiebung*) des occlusives : **bh* > *b*, **b* > *p*, **p* > *ph*. Les tenants de cette thèse s'appuyaient notamment sur les nombreux échanges Π/Φ, T/Θ, K/X observés dans les inscriptions grecques de l'aire phrygienne. Or

a) M. Lejeune (1979) a montré son extrême fragilité à partir du matériel proprement phrygien : certes **bh* aboutit bien à *b*, mais les exemples sûrs illustrent la stabilité de **d* et de **t*, cf. (αβ)βερετ(ορ) (*passim*) < **bher*, *podas* accusatif pluriel du nom du pied (G-02c), *matar* (déjà cité).

b) Les échanges graphiques du type T/Δ dans les textes grecs de Phrygie ne sont pas plus fréquents qu'à Athènes, par exemple, et, si ceux du type T/Θ y sont particulièrement abondants, c'est en raison de la rencontre, là, d'une langue possédant des aspirées (le grec) avec une autre qui n'en possédait pas (le phrygien) : le phrygophone qui apprenait le grec

21. En dernier lieu Haas, 209-212; *LB* 19 (1976) 3, 78, *sq*.

assimilait tout naturellement l'aspirée du grec à la sourde correspondante, T et Θ devenant dès lors des graphèmes interchangeables ; le même résultat s'observe partout où les mêmes conditions sont remplies : langue du Scythe des *Thesmophories* d'Aristophane ou grec d'Egypte [22].

Les labiovélaires indo-européennes ont perdu leur appendice, cf. la conjonction *ke*/κε (*passim*) ou l'indéfini κος (n° 18, 64).

On a prétendu également que le phrygien était une langue de « type *satem* » : si le second élément de l'anthroponyme composé *Benagonos* (G-116), par exemple, réfère bien au thème *$genH_1$, il montre que la langue n'a pas participé à la plus ancienne vague de palatalisation indo-européenne, celle qui fait le départ entre groupes *satem* et *centum*. Le phrygien semble bien appartenir au dernier. En réalité, les dorsales *k/g* paraissent s'être palatalisées devant voyelles antérieures *e/i* dans une phase ultérieure : cf. ζεμελως ou variantes (*supra* § 1.4) ou ζευμαν· τὴν πηγήν, Φρύγες (Hésychius), équivalent probable du grec χεῦμα (racine **ghew*). Les anciennes labiovélaires pourraient ne pas avoir été touchées (cf. indéfini κιν, n° 100), ce qui impliquerait que leur délabialisation serait postérieure à la palatalisation en question [23].

La voyelle d'appui accompagnant la réalisation des sonantes en certains contextes avait le timbre *a*, cf. *onoman* (grec ὄνομα, W-01b) ou *materan* (acc. sing. déjà cité).

Comme l'indiquent ces mêmes exemples, l'opposition */m-n/* était, comme dans la grammaire de toutes les langues indo-européennes, neutralisée en finale, avec archiphonème */N/*.

Il faut ajouter qu'en néo-phrygien la nasale finale s'efface, avec éventuelle apparition – comme dans le grec contemporain – d'une nasale parasitaire (cf. κακεν pour κακε, *supra* § 2.1.1) susceptible parfois de devenir la norme, ainsi le fréquent σεμουν (datif sing. masc.-neutre d'un démonstratif, < **se/o + me/o, passim*) en face du rare σεμου (*e.g.*, n° 4*bis*).

2.2. La morphologie

2.2.1. La morphologie nominale et pronominale

Bien que nous n'ayons identifié aucun paradigme complet, il est possible d'esquisser le système désinentiel du phrygien.

22. Voir Brixhe, 1987, 110 *sq*, et 157 (Phrygie et Asie Mineure); *L'Étranger dans le monde grec* (R. Lonis édit.), Nancy 1988, 119-123 (le Scythe); *REG* 103 (1990), 223, « Bulletin de dialectologie grecque », n° 46 (Egypte).

23. Discussion détaillée chez Brixhe, 1982, 235-246.

La flexion thématique est la mieux connue : **sing.** nom *-os/*-ος, acc. *-un/*-ου(ν), -υ(ν) [*-u*], gén. paléo-phr. *-ovo*, d'où néo-phrygien -ου [*-u*], dat. paléo-phr. *-oy* [*-ōj*] d'où néo-phr. -ου(ν) [*-u*] ; **plur.** nom. *-oi/*-οι, dat. -ωσ(ι) (pour une partie de ces finales, voir Brixhe, 1990, 94-97).

L'ouverture de **ē* en *ā* a donné à certaines déclinaisons une couleur originale, cf., pour le nom de la mère, nom. *matar*, acc. *materan*, dat. *materey*, ou encore anthroponyme Ιμαν, gén. Ιμενος dans les textes grecs d'époque gréco-romaine : l'opposition de quantité (**ē* - *e*) s'est muée en une opposition de timbre et de quantité (paléo-phr.), puis simplement de timbre (néo-phr.).

Ce que l'on comprend des documents jette quelques lueurs sur la déixis et l'anaphore. Les thèmes **so/*to*, « primitivement » complémentaires, ont été dédoublés : **to* est affecté à l'anaphore, **so* à la déixis, cf., pour le premier, la corrélation ιος νι ..., τος νι ... (n° 6, 25 ; « quiconque ..., celui-là ... ») ou le génitif *tovo*, d'où néo-phr. του (dans lequel se confondent génitif après contraction et datif < **tōj*) ; pour le second, nom./acc. neutre sing. *si=hoc*, gén. et dat. fém. sing. σας et σα(ι), et le remarquable dat. sing. masc. et neutre σεμουν ou variantes (< **se/o* + *me/o*).

Le thème **swe* pourrait avoir, lui aussi, fourni un anaphorique, cf. dat. οι.

Le phrygien a connu, comme le grec, un thème *auto-*, exprimant l'identité et l'ipséité (cf. αυτος, n° 12), susceptible de se combiner avec le précédent pour constituer un réfléchi (?), cf. *venavtun* (W-01b) et εαυται (n° 116, Brixhe-Neumann, 1.12, et p. 182).

Enfin, son relatif est *yos*/ιος (*passim*) et son indéfini est connu par le masc. κος et le neutre κιν [24].

2.2.2. La morphologie verbale

Dans le système verbal, on discerne tout d'abord une opposition de voix, entre actif et médio-passif : αδδακετ – αδδακετορ (< **dhH*$_1$- ou **dheH*$_1$-), αββερετ – αββερετορ (**bher-*), αββιρετο et αββερετοι pouvant correspondre à une erreur de lecture (le premier) ou à une faute du graveur (l'un et l'autre), cf. Brixhe, 1979, 177 *sq.*

On y a repéré :

a) un thème de présent (cf. les formes précédentes) ;

b) peut-être un thème de futur en *-s-* (*egeseti* P-04a / εγεσιτ n° 58, v. Brixhe-Lejeune, 238) ;

c) à coup sûr un prétérit à augment avec 3e pers. du sing. marquée par *-s*, cf. *edaes*/εδαες (=*fēcit*, < **dheH*$_1$-) ou les composés, *eneparkes*/

24. Pour l'ensemble des pronoms, voir Brixhe, 1978/1, 6-22, et, pour σεμουν, 1990, 95.

ενεπαρκες (Brixhe-Lejeune, 14), ποσ-εκανες (n° 116 = Brixhe-Neumann, 1.7, et p. 175-176);

d) un thème de parfait caractérisé par le redoublement, cf. τετικμενος (*passim*) ou γεγαριτμενος (n° 33, 36, 79).

Outre l'indicatif, on a quelques raisons de croire à l'existence – à côté d'un subjonctif (αδδακετ, αββερετ ?) – d'un optatif, cf. *kakoioi* et *kakuioi* (G-02c, P-04b) d'un dénominatif de *kako-*, «rendre» ou «devenir kako-».

De l'impératif, nous n'avons que des 3[e] personnes; à l'actif, sing. *-to*/-του (< *-tōd*, *si↑eto*, ειτου), plur. -νου/-τ(τ)νου; au moyen, sing. *-do*/-δου (< **dhō* ? *lakedo*, εγεδου)[25].

Enfin, bien que l'étymologie des radicaux concernés soit discutée, on a découvert plus haut, à l'occasion de l'évocation du parfait, l'existence d'un participe médio-passif en *-meno-*.

2.3. L'organisation de la phrase

La relative minceur du corpus, la quasi-obscurité des textes de quelque longueur et le caractère stéréotypé de ceux que l'on comprend à peu près font que la syntaxe est, en l'état actuel du déchiffrement, la parente pauvre.

Les imprécations néo-phrygiennes permettent d'isoler une corrélation du type ιος νι (relatif + particule à valeur indéfinie)..., Ø..., avec diverses variantes : ιος νι..., τος...; ιος νι..., ιος; etc., structure qui devrait se retrouver en quelques textes paléo-phrygiens; la subordonnée relative peut y être remplacée par une hypothétique qu'introduit αι κος «si quelqu'un..» ou αινι κος «ou si quelqu'un...» (Brixhe, 1978, 15-21).

On voit par là que la conjonction hypothétique est fournie, comme en grec, par l'ancien thème pronominal non personnel **e/o*.

Le même αινι que nous venons d'apercevoir peut aussi jouer le rôle de conjonction disjonctive «ou» : comme le latin *sive*, αινι vaut donc «ou»/«ou si». Nous avons entrevu que la conjonction copulative normale est *ke*/κε, susceptible d'être relayée par ακκε, qui pourrait avoir la même origine que le latin *atque* (Brixhe, 1978, 1, 1-6).

2.4. Lexique et onomastique

Au cours des développements précédents, nous avons aperçu certaines des unités lexicales identifiées en phrygien. On reviendra plus loin sur

25. Se reporter à Brixhe, 1979, 177-184, et 1990, 90-91.

quelques-unes d'entre elles. Signalons la présence de quelques emprunts : au grec, cf. *soroi*/σορου (*supra* § 2.1.1), θαλαμει (n° 4), κορου (grec χῶρος ? n° 92) ; au galate, peut-être, cf. τοτος/τευτους (indo-européen **teutā* « peuple, cité » ? v. Brixhe, 1979, 186-187).

Les inscriptions paléo-phrygiennes et surtout les graffites de Gordion fournissent un certain nombre de noms de personnes (voir index de Brixhe-Lejeune). Les documents néo-phrygiens, en raison de leur nature, en sont quasiment dépourvus. En revanche, les inscriptions grecques contemporaines de l'aire phrygienne sont, de ce point de vue, d'une extrême richesse (voir Zgusta, 552-558).

Du paléo- au néo-phrygien, l'anthroponymie semble être restée stable ; tout au plus avons-nous l'impression que les noms d'origine hittito-louvite sont devenus un peu plus fréquents. Que rencontrons-nous ?

– Des noms proprement phrygiens :

a) *Iman*/Ιμαν – Ιμενος (issu sans doute d'un appellatif), *Aladis, Voine(s), Mane*, Ουεναουια/Ουεναυια, Ξευνα/Ξευνη/Ξευνε, etc. ;

b) *Imeneia, Dumasteia, Lagineios*, etc., qui sont clairement des dérivés ;

c) quelques formes comme *Benagonos*, sûrement un composé.

– Des noms d'origine hittito-louvite tels que *Mamutas*, Τουης, Τροκονδας ;

– Surtout, comme partout ailleurs en Anatolie, des *Lallnamen* : *Ata(s)*, Γα, Δα, Να/Νη, *Mama, Nana*/Νανα, *Nina, Baba/B(b)a*/Βαβα(ς), etc.

2.5. L'extinction du phrygien

Les sources antiques nous parlent du phrygien pour la dernière fois au V[e] siècle (cf. Friedrich, 1941, col. 868-869) ; mais il est fort possible qu'il n'ait pas été réellement éliminé avant les invasions arabes du VII[e] siècle (Brixhe, 1987, 11).

3. Le phrygien et les langues anatoliennes

À leur arrivée en Asie Mineure, les Phrygiens ont recouvert des peuples de langue anatolienne, sans doute de type hittite au Nord et au Nord-Est de leur domaine, de type louvite au Sud et au Sud-Est. Ce contact

s'est traduit, on l'a vu, par l'introduction de certains noms hittito-louvites dans l'anthroponymie phrygienne.

Il a eu une autre conséquence, sur la grammaire, celle-là : les langues de l'adstrat (puis du substrat) ignoraient l'opposition morphologique masculin-féminin et face au neutre (non animé) elles ne connaissaient qu'un genre commun (animé). Elles vont engendrer en phrygien un flottement dans la finale du nominatif singulier des anthroponymes masculins (d'où *Atas/Ata*, *Voines/Voine*) et féminins (d'où Ναvα/Ναvας), voir Brixhe, 1983, 128. On ne sait si cette hésitation, qui se retrouve dans toute l'anthroponymie anatolienne des époques hellénistique et romaine (Brixhe, 1987, 77-79) a dépassé le cadre des noms de personnes.

4. Le phrygien parmi les langues indo-européennes

On a déjà vu que le phrygien était sans doute à mettre au nombre des langues *centum*, quelle que soit l'appréciation portée sur ce trait.

On écartera, en outre, l'idée, répandue par Kretschmer, d'une unité thraco-phrygienne. Le thraco-dace ou le thrace et le daco-mésien semblent appartenir au groupe *satem* et sa/leur phonétique est, d'après ce qu'on entrevoit, beaucoup moins conservatrice que celle du phrygien (voir *Le Thrace*, §§ 3 *sq.*).

Le phrygien partage avec la famille hittito-louvite quelques isoglosses (*e.g.* désinence *-s* de 3e pers. du sing. du prétérit, *r* marque médio-passive, Brixhe, 1968, 316-317) ; mais aucune n'est absolument nette, aucune non plus n'est exclusive.

Les concordances gréco-phrygiennes sont infiniment plus éclairantes. Isoglosses non exclusives : le pronom relatif *yos*, l'utilisation de l'augment. Isoglosses propres aux deux langues : la marque *-s* du nominatif sing. des masculins en *-ā*, le suffixe participial *-meno-*, le pronom *auto-* et le thème *kako-*, qui ne sauraient être des emprunts au grec puisque déjà présents en paléo-phrygien, *ai* enfin, qui connaît le même emploi hypothétique que le αἰ dorien et éolien. Ces traits trahissent de toute évidence d'étroits rapports préhistoriques entre le grec et le phrygien : simple voisinage géographique ? plutôt appartenance à un même groupe ? Les recherches futures confirmeront ou infirmeront l'hypothèse. Cette parenté pourrait être illustrée par la titulature de Midas, sur le célèbre « Monument de Midas » (M-01 a), *Midai lavagtaei vanaktei* (datif) : les deux fonctions ici évoquées, qu'on retrouve en mycénien (*lawagetas* et *wanax*), ont toutes chances non d'avoir été empruntées, comme on le croit généralement, mais

d'être autochtones et, par conséquent, d'être le fruit d'un héritage commun (Brixhe, 1990, 73-75).

On peut, en outre, songer à des rapports – plus anciens sans doute – avec les ancêtres des Latins. C'est du moins ce à quoi semblent nous inviter quelques concordances intéressantes : utilisation du préverbe *ad* (αδ-δακετ, αβ-βερετ), extension de l'infixe *-k-* au thème du présent (αδδακετ = probablement *afficiat*), *-tor* désinence de la 3[e] personne du singulier médio-passive (= lat. *-tur*), ακκε peut-être, s'il équivaut au lat. *atque*.

Bibliographie

BRIXHE Cl. (1968). – *RPh*, 42, 306-319 (c.r. de Haas 1966).

– (1975). – «Problèmes d'interprétation du phrygien», *Le Déchiffrement des écritures et des langues* (Colloque du XXIX[e] Congrès des Orientalistes, Paris, 1973, présenté par J. Leclant), Paris, 65-74.

– (1978). – «Etudes néo-phrygiennes», I, *Verbum* I 1, 3-21.

– (1978/1). – «Etudes néo-phrygiennes» II, *Verbum* I 2, 1-22.

– (1979). – «Etudes néo-phrygiennes» III, *Verbum* II 2, 177-192.

– (1982). – «Palatalisations en grec et en phrygien», *BSL* 77, 209-249.

– (1983). – «Epigraphie et grammaire du phrygien : état présent et perspectives», *Le lingue indoeuropee di frammentaria attestazione* (Atti del Convegno della Società Italiana di Glottologia e della Indogermanische Gesellschaft, Udine, 1981), Pise, 109-133.

– (1986). – «Les Grecs, les Phrygiens et l'alphabet», *Mélanges Mihailov*, Sofia (sous presse).

– (1987). – *Essai sur le grec anatolien au début de notre ère*[2], Nancy.

– (1987/1). – «Les inscriptions paléo-phrygiennes de Tyane : leur intérêt linguistique et historique», in *La Cappadoce méridionale jusqu'à la fin de l'époque romaine. Etat des recherches* (Colloque Istanbul 1987), ed. par Br. Le Guen-Pollet et O. Pelon, Paris 1991, 37-46 et pl. I.

– (1990). – «Comparaison et langues faiblement documentées : l'exemple du phrygien et de ses voyelles longues», *La Reconstruction des laryngales*, Liège, 59-99.

BRIXHE Cl., LEJEUNE M. (1984). – *Corpus des inscriptions paléo-phrygiennes*, Paris.

BRIXHE Cl., NEUMANN G. (1985). – «Découverte du plus long texte néo-phrygien : l'inscription de Gezler Köyü», *Kadmos* 24, 161-184.

BRIXHE Cl., WAELKENS M. (1981). – «Un nouveau document néo-phrygien au Musée d'Afyon», *Kadmos* 20, 68-75.

DIAKONOFF I., NEROZNAK V.P. (1985). – *Phrygian*, New York.

FRIEDRICH J. (1932). – *Kleinasiatische Sprachdenkmäler*, Berlin.

– (1941). – *RE* XX-1, col. 868-891, *s.v. Phrygia : Sprache* et *Phrygia : Geschichte*.

HAAS O. (1966). – *Die phrygischen Sprachdenkmäler (=LB* 10), Sofia.

LEJEUNE M. (1979). – «Regards sur les sonores indo-européennes en vieux phrygien», *Mélanges Laroche*, 219-224.

NEROZNAK V.P. (1978). – *Paleobalkanskije jazyki*, Moscou.

NEUMANN G. (1988). – *Phrygisch und Griechisch* (Sitzungsb. Österr. Akad. Wiss., Philos.-hist. Klasse, 499), Vienne.

ZGUSTA L. (1964). – *Kleinasiatische Personennamen*, Prague.

CHAPITRE 9

LE THRACE[1]

Claude BRIXHE, Anna PANAYOTOU

Ce n'est guère que vers le milieu du XIX[e] siècle qu'on commença à porter un intérêt particulier à la ou aux langues parlées jadis dans l'Est de la péninsule des Balkans. Le véritable fondateur de la thracologie fut W. Tomaschek et, jusqu'à la fin de la Seconde Guerre, les études portant sur cette aire furent essentiellement concentrées à Vienne. Après la guerre, parallèlement à une intense activité archéologique, elles se déplacèrent vers l'Est, vers Belgrade et Skopje, mais surtout vers Bucarest et Sofia, où fut fondé en 1972 l'Institut de thracologie (langue, histoire, archéologie).

De nombreux développements ont été consacrés ici ou là à l'histoire des recherches linguistiques [2]. On se contentera ici de mentionner le nom de

1. Ce titre ne constitue naturellement pas une prise de position en faveur d'une des thèses qui seront évoquées plus loin. Il représente une désignation commode pour la ou les langues de l'espace visé par la thracologie. – La plupart des formes utilisables sont fournies par des textes grecs, littéraires ou épigraphiques : dans les pages qui suivent, aucune forme, quelle qu'en soit l'origine, ne sera accentuée ; ce nous semble être la solution linguistiquement la plus satisfaisante.

2. Voir, par exemple, Russu, 32-48, 149 [63-89] ; Georgiev, 1983, 1149-1150 ; *id.*, 1984, 207-208 ; Velkova, 13-38.

ceux qui en ont marqué les étapes les plus importantes : outre Tomaschek (1893 et 1894), P. Kretschmer, qui en 1896 lança la fameuse thèse thraco-phrygienne ; D. Detschew, qui en 1957 publia le plus complet des recueils de données linguistiques thraces, imparfait, constamment amélioré, mais irremplaçable ; enfin I.I. Russu (1959, 1969) et Vl. Georgiev (1957, 1966, 1983...), qui dans les trente dernières années ont animé des débats souvent ponctués par les interventions mesurées de G. Mihailov (1984, 1986, 1987).

1. L'ESPACE ET LE TEMPS

Le premier obstacle rencontré réside dans la définition de l'aire thrace. Certes, elle a parfois été très vaste, fluctuant énormément de la préhistoire à l'histoire ; mais, dans les solutions avancées affleure parfois une pointe de chauvinisme, un peu semblable à celui qui, aux XVIIe et XVIIIe siècles, faisait exalter les Gètes et les Daces par l'historiographie saxonne de Transylvanie, à partir de l'identification des Gètes et des Goths, des Daces et des Detschen/Deutschen [3] : d'assimilation en assimilation, J.C. Dragan en arrive à annexer à l'un de ses articles une carte où « l'espace thrace » va de la Baltique à la Méditerranée, du Dniepr et du centre de l'Anatolie au Sud de l'Espagne, englobant la Grèce continentale (sauf le Péloponnèse) et l'Italie (y compris la Sicile) [4].

Ne retrouverait-on pas, en plus discret, le même parti pris dans certaines tentatives d'ethnogenèse des Thraces et des Daces, sur bases linguistiques et archéologiques ? On affirme ici [5] que l'actuelle Roumanie, habitat primitif des Géto-Daces, fut le centre principal de la formation et de la diffusion de l'indo-européen commun au Nord et au Sud du Danube. On prétend là [6] que le Sud de la Bulgarie, c'est-à-dire la « Thrace proprement dite » est occupée, sans solution de continuité, depuis le néolithique (dès le Ve et même le VIIe millénaire avant J.-C.) par les Thraces, qui en furent les premiers habitants. De telles affirmations ne peuvent que laisser pour le moins sceptique, quand on sait comment se font les ethnies et les langues et quand on connaît la perpétuelle instabilité (historique) de la région.

---

3. Cf. Gh. Bulgăr, *Troisième Symposium*, 239-240.

4. *Deuxième Symposium, in fine*. On en trouvera l'explication sous la plume du même *Premier Symposium*, 36. Notons que J.C. Dragan est l'auteur de *We, the Thracians*, Milan, 1976, qui semble avoir été réédité en français et italien.

5. Ainsi, A. Vraciu, *Thraco-dacica*, 320-321.

6. Cf. Vl. Georgiev, 1966, 170-171, et G.I. Georgiev, *Dritter Kongress* I, 67-74.

La thèse traditionnelle, depuis Tomaschek et Kretschmer, fait arriver les Thraces dans les Balkans au début du IIe millénaire, avec la grande migration des peuples indo-européens.

1.1.

Pour les Anciens la Thrace allait de l'Egée au Danube, excluant les territoires conquis par les Macédoniens (*infra* § 1.2.1), les colonies grecques (sur l'Egée et la mer Noire) et l'enclave péonienne sur les cours moyen et supérieur de l'Axios et entre Strymon et Nestos.

Cette définition même implique une extension antérieurement plus large des tribus thraces.

1.1.1.

On voit celles-ci intervenir dans le passé mythique de la Phocide (Thucydide, II 29.3; Strabon, VII 7.1 et IX 3.13), de la Béotie (Strabon, IX 2.25) et même de l'Attique (*id.* VII.7.1). Le même Strabon nous rapporte (*ibid.*) que de son temps il y avait encore des Thraces en Thessalie, sans que ces indications reçoivent le soutien d'autres témoignages.

On a voulu identifier des anthroponymes thraces dans les documents en linéaire B et même en linéaire A [7]. Si les identifications étaient exactes, elles révéleraient des contacts entre Thraces et Minoens ou Grecs achéens, sans nous dire où étaient alors les Thraces.

La présence de Thraces est suggérée, à Thasos, par le mythe et le nom ancien de l'île selon Hésychius, Ἠδωνίς, qui rappelle celui d'une tribu thrace localisée en face, sur le continent (les Ἠδῶνες). A Lemnos, les plus anciens habitants connus, les Σίντιες/Σιντοί (Homère, *Il.* I 593-594; *Od.* VIII 294), seraient, selon Strabon (VII, frag. 45 et 45a) d'origine thrace [8]. Deux des noms anciens de Samothrace, Σαόννησος et Σαωκίς, sont rapprochés du nom d'une tribu thrace, les Σάϊοι, en face, sur le continent [9]. Ces indications ne sont confirmées actuellement que pour Samothrace, par la présence de documents écrits non grecs; mais, même si la célèbre stèle de Lemnos n'est pas thrace – ce qu'après tout personne ne peut prouver – il y a quelques chances pour que les Thraces aient, avant la colonisation grecque, occupé les îles du Nord de l'Egée.

1.1.2.

De nombreux témoignages anciens attestent une importante migration thrace dans le Nord-Ouest de l'Asie Mineure (rive asiatique de la

7. Cf. P.G. van Soesbergen, *Ancient Bulgaria* I, 199-212; I. Duridanov, *Contributions*, 110-117; J. Best, *Thracians and Mycenaeans*, 135-142.

8. Cf. les Thraces Σιντοί, au Sud du pays péonien, Thucydide, II 98.1-2.

9. *IG* XII 8, p. 36-37.

Propontide, Mysie, Bithynie), intervenue vraisemblablement à la fin du IIe millénaire avant J.-C. Ainsi, par exemple, Hérodote (I 28) parle des Θρήικες οἱ Θυνοί τε καὶ Βιθυνοί (cf. encore VII 75, à propos des Bithyniens). En VI 34, il nous apprend que «εἶχον Δόλογκοι Θρήικες τὴν χερσόνησον ταύτην (de Cyzique)». En III 90, il évoque les Θρήικες οἱ ἐν τῇ Ἀσίῃ. Voir encore le développement plus complet de Strabon (XII 3.3), qui, outre la tribu des Βέβρυκες installée en Mysie, fait venir de Thrace les Mysiens (οἱ Μυσοὶ Θρᾳκῶν ἄποικοί εἰσι τῶν νῦν λεγομένων Μοισῶν). Il y a, en effet, certainement identité d'origine entre les Mésiens (Μοισοί) d'Europe et les Mysiens (Μυσοί) d'Asie et, contre Hérodote (VII 20 et 75) qui suggère un passage d'Asie en Europe, Strabon et d'autres ont probablement raison d'envisager le processus inverse. Les deux groupes de tribus mysiennes, d'ailleurs souvent désignés par le même nom, Μυσοί (la forme ancienne), ont fini par recevoir deux appellations différentes (Μυσοί pour les Asiatiques, Μοισοί pour les Européens)[10], sans doute à l'époque où Y et OI recouvraient en grec tous deux le son [*y*][11].

Ces témoignages sont confirmés par l'épigraphie : d'après l'onomastique personnelle des époques hellénistique et romaine, les Thraces semblent constituer un élément important de la population à Cyzique, Kios, Nicée, Prusa, Nicomédie, etc[12]; en Troade, ils paraissent avoir été plus clairsemés. Entre les deux rives de la Propontide, à époque historique, le flux des migrants thraces dans les deux sens pourrait avoir été constant[13].

En revanche, quand aux mêmes époques on rencontre des Thraces plus à l'intérieur, en Lydie, en Carie ou en Pisidie, il s'agit vraisemblablement d'une immigration récente : mercenaires amenés par les Attalides ou les Séleucides[14]. De même, c'est l'importance des contingents thraces dans les armées lagides qui justifie l'abondance des noms thraces en Egypte[15].

---

10. D'où latins *Mysi, Moesi* : *Moesi* a pu même être parfois utilisé pour les deux fractions, cf. Solinus, 40.20 (milieu du IIIe siècle ap. J.-C.).

11. *Contre* (évolution phonétique proprement «daco-mésienne»), Georgiev 1966, 159 (sans argument sérieux). Le même auteur, moins sûr de son hypothèse, évoque ailleurs les deux explications, *Premier Symposium*, 24, Sur l'ensemble du problème «Mésien/Mysien», on pourra consulter utilement Papazoglou, 391-437, qui rassemble, 587-600 (M1 à 55) tous les témoignages antiques.

12. Pour s'en persuader, il suffit de consulter les index des corpus concernés.

13. Sur cette présence en Mysie, mais surtout en Bithynie et sur la Propontide, voir notamment St. Mitchell, *Pulpudeva* 2, 119-127. Pour Cyzique se reporter, par exemple, à E.S. Goloubtsova, *Dritter Kongress* II, 236-239.

14. Cf. St. Mitchell, *o.c.*, 122, ou L. Robert, *Villes d'Asie Mineure*[2], Paris, 1962, 33, 235-236, 250, 414-415.

15. Voir, par exemple, M. Launey, *Recherches sur les armées hellénistiques*[2] II, Paris 1987, 1191-1203; V. Velkov et A. Fol, *Les Thraces en Egypte gréco-romaine* (=Studia Thracica 4), Sofia 1977; et J. Bingen, *Pulpudeva* 4, 72-79.

1.2.

Donc, au premier millénaire avant J.-C. au moins, on a parlé le thrace dans une bonne partie du Sud-Est de la péninsule des Balkans et au Nord-Ouest de l'Asie Mineure.

Peut-on préciser les contours de ce périmètre et en apprécier l'homogénéité ethnique ?

1.2.1.

Limites Nord : Pour l'Antiquité classique, la Thrace allait jusqu'au Danube. A vrai dire, elle ne savait pas grand-chose sur la situation ethnico-linguistique au Nord de ce fleuve. Hérodote nous dit que, sur la mer, donc au Nord-Est, la Thrace touchait à la Scythie (IV 99), mais qu'au Nord s'étendait une zone déserte illimitée (V 9). On sait pourtant que les tribus daco-gètes allaient jusqu'au cœur des Carpates [16] et que le Danube ne constituait pas une frontière infranchissable, les Gètes (selon Strabon, VII 3.13) passant perpétuellement le fleuve pour se mêler « aux Thraces et aux Mésiens ».

Au Sud-Ouest, on rencontre un nombre relativement faible de noms thraces en Macédoine occidentale [17], mais davantage en Macédoine centrale et orientale [18]. C'est que depuis l'époque archaïque au moins, avec la constitution du royaume, les Macédoniens empiètent constamment sur la Thrace. Pour Hécatée, celle-ci commençait encore à l'Est de l'Axios ; avec Alexandre Ier, la Macédoine atteint le Strymon. Poussée d'Ouest en Est, donc. Des populations émigrent, laissant peut-être cependant sur place des îlots : ainsi les Pières, d'abord autour de l'Olympe selon Strabon (VII frag. 11, et IX 2.25), en sont délogés et se retrouvent au pied du Pangée, à l'Est du Strymon, au temps d'Hérodote (VII 112 ; cf. Thucydide, II 99.3). Les Edoniens, dont les Mygdoniens, jadis entre l'Axios et le Strymon, en sont chassés (Thucydide, II 99.4) et Xerxès les rencontre déjà à l'Est du Strymon (Hérodote, VII 110). D'autres tribus voient leur territoire se rétrécir, comme les Bisaltes, qui, après avoir, semble-t-il, occupé un large espace [19], sont restreints à la péninsule orientale de la Chalcidique et à la rive Ouest du Strymon (Thucydide, IV 109.4).

Après les migrations du deuxième millénaire avant J.-C., les Thraces ont donc occupé une aire très vaste, de la mer Noire à l'Ouest de l'Axios,

---

16. Sur la frontière septentrionale du monde thrace, voir M. Dušek, *Pulpudeva* 1, 108-131.

17. Cf. J. Touratsoglou, *Pulpudeva* 2, 128-146. Naturellement, la faiblesse apparente de l'implantation et la date des documents (I[er] siècle avant J.-C.-IV[e] siècle après J.-C.) empêchent de savoir s'il s'agit d'un résidu ancien ou d'un apport récent.

18. Voir *e.g.* G. Bakalakis, « Thrakische Eigennamen aus den nordägäischen Küsten », *Thracia* II (Primus Congressus Studiorum Thracicorum), Sofia 1974, 261-279.

19. Cf. D. Samsaris, *Ἱστορικὴ γεωγραφία τῆς Ἀνατολικῆς Μακεδονίας κατὰ τὴν ἀρχαιότητα*, Salonique 1976, 54-56.

aire qui au Sud-Ouest n'a cessé de se réduire depuis l'époque archaïque, sous la pression macédonienne, pour être finalement limitée par le Strymon depuis Alexandre Ier et par le Nestos avec Philippe II.

1.2.2.

Pour pouvoir apprécier les matériaux recueillis, le linguiste a besoin de délimiter le périmètre concerné, mais aussi d'évaluer l'homogénéité des populations qui y vivaient.

I.I. Russu, 18 [27], dresse une liste de 104 tribus [20], incluant quelques unités micrasiatiques et quelques autres d'existence douteuse, refusant de faire une place particulière aux Daces et aux Gètes. Il donne en fin de volume(s) une carte, qui localise les principales : Odryses, Besses, Triballes, etc. [21]. Il semble que primitivement Θρᾷκες/Θρήικες ait désigné une tribu du Sud-Est de la Péninsule (sur l'Hébros), dont le nom a été étendu par les Grecs à toutes les autres [22].

Rompent cet ensemble les enclaves péoniennes sur les cours moyen et supérieur du Nestos et de l'Axios. Mais doit-on avec Russu et d'autres considérer le reste comme appartenant à un même peuple, ayant même langue avec éventuelle diversification dialectale ? Les Dardaniens, par exemple, sur le cours supérieur de l'Axios et de la Morava yougoslave, sont, comme le veut la tradition antique, tenus pour illyriens par Papazoglou [23], mais pour apparentés aux Mésiens et aux Daces par Georgiev [24].

Surtout, Georgiev (1957, etc.) a-t-il raison de refuser la qualité de thrace aux Daco-gètes ? « Dace » et « Gète » sont, nous dit Strabon (VII 3.12), deux appellations référant à un même groupe de tribus, « Gète » s'appliquant à celles de l'Est de part et d'autre du Danube, « Dace » à celles de l'Ouest. Gètes et Daces, qui parlaient la même langue (VII 3.13), étaient considérés par les Grecs comme thraces, selon le même Strabon (VII 3.2), cf. encore Hérodote, IV 93, 95-96, 118. D'où le refus de Russu et d'autres de les séparer de l'ensemble thrace [25]. C'est sur les critères linguistiques qu'on verra plus loin que Georgiev fait le départ entre les Thraces proprement dits au Sud du flanc septentrional de l'Haimos (les Balkans) et le groupe daco-mésien, totalement différent du point de vue

---

20. Il y en avait probablement davantage, mais il est difficile de faire le départ entre nom tribal et ethnique (lié à une localité, par exemple).

21. Sur ce point encore, les données ne sont pas toujours sûres : voir les listes restreintes, avec tentative de localisation, fournies par Katičić, 134-135; Georgiev, 1983, 1151, et 1984, 208.

22. Cf., par exemple, Russu, 16-17 [24-25]; Georgiev, 1984, *l.c.*; Velkova, 7-8.

23. 210-218 et 551-571 (*testimonia*).

24. 1966, 156; *Premier Symposium*, 24.

25. Sur leur histoire, voir Gh. Bichir, *Thraco-dacica*, 287-307.

ethnique, avec Daco-Gètes, Mésiens, Triballes, etc.[26], au Nord du Danube (la Dacie romaine) et au Nord-Ouest de la Péninsule (la Mésie Supérieure). Entre les monts Balkans et le Danube, une zone intermédiaire (la Mésie Inférieure), primitivement occupée par les Thraces, puis envahie par les Daco-Gètes. Cette hypothèse est aujourd'hui largement répandue.

1.2.3.

Si l'on ajoute qu'en dehors des intermèdes odryse (V^e^-IVe siècle av. J.-C. pour l'actuelle Bulgarie) et gète (Burébistas, Ier siècle av. J.-C., au Nord, mais aussi au Sud du Danube) cette région n'a jamais connu d'unité que sous tutelle étrangère, qu'elle a très tôt subi la pression des Grecs, puis des Macédoniens, puis des Romains, des Celtes, des Germains et enfin des Slaves, on devinera aisément, à la lumière des développements précédents, la complexité des problèmes proprement linguistiques.

2. La documentation

Pour approcher la langue de l'ensemble thraco-dace et ses éventuels dialectes (Russu), ou les langues thrace et daco-mésienne et leurs éventuels dialectes (Georgiev), de quels documents disposons-nous ?

2.1. Les documents en langue thrace

2.1.1.

On écartera tout d'abord de ce dossier un certain nombre de textes :

a. L'inscription « mysienne » d'Üyücek (Haut Rhyndacos, aux confins de la Phrygie, de la Lydie et de la Mysie, Friedrich, 140-141, XI, IVe-IIIe siècles av. J.-C.). Certains thracologues l'utilisent comme illustrant la langue des Mysiens d'Asie[27] et même l'incluent dans leur corpus[28]. Or elle appartient à un autre ensemble, cf. « Phrygien », § 1.3.

26. Près de 150 noms tribaux ou ethniques connus pour les Thraces, une cinquantaine pour les Daces, cf. C. Poghirc, 1976, 340.

27. Ainsi, Georgiev, 1966, 144-145, et 1983, 1192 ; V. Pisani, *Deuxième Symposium*, 158 ; Katičić, 150 (chez ces trois auteurs, avec lecture, l.7, *patrizi* à abandonner pour *patriji*).

28. Cf. Neroznak, 36-37, B1 (lecture *patriji*, l.7), qui lui fait une place à part entre textes et gloses thraces.

b. Les « propos » du dieu triballe des *Oiseaux* d'Aristophane, dont il n'y a rien à tirer pour le thrace [29].

c. L'inscription prétendument dace trouvée près de Dialul Grîdiştiei en Roumanie : sur un vase d'argile, donné par deux sceaux, en caractères latins, *Decebalus Perscorilo*, segmenté *per Scorilo* par Georgiev [30], qui voit dans *per* l'avatar de **pwero* (lat. *puer*), suivi d'un ablatif marquant l'origine « fils de Scoril ». Qu'on le lise *Decebalus per Scorilo* ou *Perscorilo*, le texte a toutes chances d'être latin [31].

d. D'authentiques textes grecs : chez Neroznak, 33-35, les numéros A 6 (῾Ιππόμαχ(ο)ς), A 7 (Κοας ΚΟΜΟΣ = *e.g.* Κōμος), A 5 (Σκυθοδοκō), A 8 (Κοτυος ἐγ Γηιστōν) [32] n'ont rien à faire ici ; quelle que soit l'origine de l'anthroponyme ou du toponyme concerné, la grammaire est grecque [33].

c et **d** sont à placer simplement aux côtés des témoignages évoqués *infra* § 2.2.2.

2.1.2.

On a une série d'inscriptions en caractères grecs provenant de la région de Plovdiv (Πιλιππόπολις/*Pulpudeva*).

a. Sur le châton d'une bague en or trouvée en 1912 dans un tumulus funéraire près du village d'Ezerovo (non loin de Părvomaj), huit lignes d'écriture en *scriptio continua* (première moitié du Ve siècle av. J.-C.), Friedrich, 148, XV ; Detschew 1957, 566-582, avec photo et les diverses tentatives d'interprétation alors apparues ; Russu, 25-27, avec dessin et bibliographie [46-49, avec photo et dessin] ; Georgiev 1966, 130-131 ; *id.* 1983, 1159-1160, avec photo, pl. 1 (entre p. 1160 et 1161) ; R. Schmitt-Brandt, *Glotta* 45 (1967), 47-51 ; Neroznak, 26-28, A 1, avec photos (après p. 112) ; etc. Nombreuses divergences dans la segmentation, d'où une bonne vingtaine d'interprétations et absence de consensus.

b. Sur un anneau d'or trouvé à la main gauche d'un squelette dans un tombeau au village de Duvanlii : autour de la représentation d'un cavalier, un texte mutilé de la seconde moitié du Ve siècle av. J.-C. : ΗΥΣΙ̣Η̣ ... Δ̣Ε̣Λ̣Ε̣ ΜΕΖΗΝΑ̣Ι (ou -ΝΔ̣Ι ; voir notamment V. Pisani, *Paideia* 16

29. Voir en dernier lieu, Cl. Brixhe, in *L'Etranger dans le monde grec*, R. Lonis édit., Nancy 1988, 114-116.

30. 1966, 144, et 1983, 1181.

31. Ainsi G. Mihailov, 1984, 209 ; Neroznak l'inclut dans son corpus (36, A10) avec lecture *Decebalus Perscorilo*, deux nominatifs selon lui : alors grammaire latine ?

32. Totalement incompris par Neroznak, voir son commentaire.

33. Voir maintenant les textes gravés sur récipients d'argent du trésor de Rogozen (Nord-Ouest de la Bulgarie), Mihailov, 1987, 5-19, avec un excellent commentaire, qui montre bien qu'il faut voir dans Σατοκō ou Διδυκαιμō (seconde moitié du Ve, début du IVe siècle av. J.-C.) des génitifs attiques. Le n° A 9 de Neroznak est tout aussi grec, sauf peut-être (?) dans sa dernière partie (παδρυ ΙΗ | Β), écrite d'une autre main, voir Mihailov, *ibid.*, 18, n. 8.

(1961), 245 ; Georgiev 1966, 131-132 (où lecture -NAI), et 1983, 1160-1161 (où lecture -NΔI) ; Neroznak, 32-33, A 4 ; en dernier lieu Vl. E. Orel, *Glotta* 64 (1986), 48-49. Autant d'interprétations que d'auteurs.

c. Sur quatre vases d'argent trouvés dans un tombeau près du même village, ΔΑΔΑΛΕΜΕ, fin du Ve – début du IVe siècle av. J.-C. ; cf. Detschew, 1957, 110 ; Georgiev, 1966, 132, et 1983, 1161 ; Neroznak, 32, A 3 (avec photo entre p. 112 et 113) ; en dernier lieu, Orel, *l.c.* : quatre interprétations différentes.

2.1.2.1.

Près du village de Kjolmen, district de Preslav (N.-E. de la Bulgarie), on a trouvé en 1965, dans un tumulus funéraire, sur une pierre plate (deux fragments jointifs), une inscription en caractères grecs archaïques, assignable aux VIe-Ve siècles av. J.-C. Elle affecte la forme d'un demi-cercle : trois lignes, A, B sur la circonférence (à la suite l'une de l'autre, l'une sinistroverse, l'autre dextroverse) [34], et C (sinistroverse) sur le diamètre : voir V. Beševliev, *Glotta* 43 (1965), 317-322 (avec photos) ; R. Schmitt-Brandt, *O.c.*, 42-47 ; Neroznak, 29-31, A 2 (photo entre p. 112 et 113) ; Georgiev 1983, 1161-1163 ; I. Moldoveanu, *Dritter Kongress* I, 223-227 (photo, 227). Interprétation parfois délirante. Il est vrai que le document pose bien des problèmes :

a) identification et ordre de lecture des lignes,

b) identification de certains symboles ;

en outre, le lieu de découverte nous interroge quant à la langue du texte, puisque nous sommes près de la Dobroudja, à la limite du pays gète : avons-nous affaire au thrace ? au daco-gète ? au parler mixte imaginé par Georgiev pour la Mésie Inférieure ? à autre chose [35] ?

2.1.2.2.

Samothrace a pu connaître un premier peuplement thrace (*supra* § 1.1.1), dont la langue est susceptible d'avoir survécu à travers des formules liturgiques au service du culte des « Grands dieux » ou « Cabires » [36]. On a trouvé là des documents en une langue non grecque :

a. Un fragment de la partie droite d'une stèle (dix lignes, donc mutilées à gauche), assignable aux Ve-IVe siècles av. J.-C. ; cf. Lehmann 1955, 100, n° 40 (avec photo, pl. 40) ; Fraser, 120-121, n° 64 (photo, pl. XXV) ; photo également chez Neroznak, entre p. 112 et 113.

34. Le graveur a manifestement changé de position et inversé l'orientation des lettres.

35. Cf. G. Mihailov, 1984, 209, qui se demande s'il s'agit d'un texte thrace (au sens trivial du mot : M. ne suit pas la thèse de Georgiev avec deux langues et deux peuples).

36. Cf. Diodore (5.47.3), qui dit μέχρι τοῦ νῦν : l'indication renvoie-t-elle à l'époque de Diodore ou à celle de sa source ? voir Fraser, 121, et Lehmann, 1960, 8.

b. 75 graffites sur vases (ex-voto), réduits parfois à une seule lettre, cf. Lehmann 1960, 45 *sqq.*, n° 1-75, auxquels il faudrait peut-être ajouter quelques autres (*ibid*, 10, n° 11). Le plus long, le n° 1, p. 45 (photo pl. I), une ligne mutilée aux deux extrémités, comporte dix-huit lettres (= L.H. Jeffery, *The Local Scripts of Archaic Greece*², Oxford 1990, 299, 307 n° 57, 412; fac-similé pl. 57); photo chez Neroznak, entre p. 112 et 113. Le n° 2, p. 46 (photo, pl. I), est un fragment présentant cinq lignes et une sixième perpendiculaire aux autres. Les autres graffites sont du type ΔΙΝΤΟΛΕ, ΔΙ/ΕΝΤ[Ο]ΛΕ, ΔΙΝ, Δ̇Ι, Δ ou ΛΕ, ΛΙ, Λ... ; il s'agit donc manifestement, le plus souvent, d'abréviations. L'ensemble s'étale du VIe au IVe siècle av. J.-C. Tentatives d'interprétation par Bonfante, 101-109, et Lehmann, 1960, 8-19; allusion chez Georgiev, *Premier Symposium*, 32.

2.1.2.3.

Aucun des documents qui viennent d'être recensés[37] n'a reçu une interprétation susceptible d'être acceptée sans réserve.

2.2. Autres sources

2.2.1.

Les auteurs anciens ont laissé un certain nombre de gloses :

a. 80 à 90 gloses thraces, surtout fournies par Hésychius et Photius, cf. en dernier lieu Velkova, qui, parmi 90 gloses, en isole 30 (dont -διζα et -παρα, extraits de la toponymie) sûrement assignables au thrace (41-84), et 22 dont l'attribution est douteuse ou seulement probable, les autres devant être attribuées au grec, au phrygien, au macédonien, au dace, etc. (103-126).

b. Une soixantaine de gloses daces, noms de plantes donnés par Dioscoride et Pseudo-Apulée, dont certaines sont irrémédiablement corrompues ou sont d'origine grecque ou latine, cf. Detschew, 1957, 541-565; Russu, 27-30, 56-82 [52-56, 106-148]; Georgiev, 1966, 140-142, et 1983, 1179-1180; Neroznak, 55-65.

À ces gloses, il faut ajouter le mot *midne* (ablatif, = *vico* ?), livré par une dédicace latine (Rome, IIIe siècle ap. J.-C.) faite à Jupiter Optimus

---

37. En raison des incertitudes qui les entourent, nous avons écarté quelques pièces : sur gemmes (Musée de Sofia), *SEG* 31, 659-661, 664, 669, les suites incompréhensibles de lettres correspondent-elles à un «texte» thrace ou à des séquences magiques sans signification? La mystérieuse inscription sur gargouille trouvée en Bisaltie et publiée naguère avec fac-similé seulement par G.B. Kaphtantzis, *Ἰστορία τῆς πῆς πόλεως Σερρῶν καὶ τῆς περιφερείας της* I, Athènes, 1967, 306, n° 507, est actuellement introuvable.

Maximus par un certain *Aur(elius) Bitus sacerdos* et des *cives Prov(inciae) Tracie Serdicens(is) : midne Potelense salvo, CIL* VI 1, 2819; voir en dernier lieu Velkova, 74-75.

2.2.2.

Les monnaies à légendes grecques et les inscriptions grecques[38] depuis le Ve siècle av. J.-C., les incriptions latines plus tard, les auteurs grecs et latins fournissent un grand nombre d'anthroponymes, de toponymes, d'hydronymes, d'oronymes et de théonymes thraces, du plus grand intérêt pour la langue[39].

2.2.3.

Enfin, doivent être également versés au dossier :

a. Les éléments prégrecs, prélatins et préslaves de la toponymie, de l'hydronymie[40] et de l'anthroponymie[41] roumaines et bulgares actuelles.

b. Les éléments prélatins et préslaves des lexiques bulgare et roumain; ainsi, sur le fonds lexical roumain (1 700/1 800 mots), près de 10% sont préromains, donc probablement daces, soit 160 à 170 mots, dont 70 pourraient se retrouver en albanais[42].

Pour la connaissance de la ou des langue(s) ancienne(s) de la région, on a tenté d'utiliser d'autres aspects des langues balkaniques modernes : phonétique roumaine, suffixation roumaine, sémantisme de certaines unités roumaines ou bulgares, etc.[43].

3. La langue

3.1. L'écriture

On a trouvé en Roumanie (1961) et en Bulgarie (1966) des vestiges de ce qui pourrait être l'écriture la plus ancienne d'Europe, puisque

---

38. Les plus anciennes ont été trouvées hors de Thrace (Athènes, Delphes). En Thrace même, les premiers documents sont assignables au troisième quart du Ve siècle av. J.-C. (monnaies, voir B.V. Head, *Historia numorum*^{2}, Oxford 1911, 282-286; trésor de Rogozen, *supra* § 2.1.1.d).

39. Ainsi Poghirc, 1976, 335-347, travaille sur près de 3 000 unités puisées à ces sources.

40. Cf. Russu, 112-113 [197-200].

41. Cf. Katičić, 152-153 (à propos de l'anthroponymie bulgare).

42. Cf. Russu, 119-131 [211-255], et *Troisième Symposium*, 339-341 ; Georgiev, 1966, 145-146; G. Bulgăr, *Premier Symposium*, 84-92; C. Poghirc, *Dritter Kongress* I, 239-244; S. Strati, *ibid.* 254-258; etc.

43. Voir C. Dumitrașcu, *Thraco-dacica*, 329-330 (*l* > *r* en roumain), ou I. Coja, *Deuxième Symposium*, 213-219 (le suffixe *-ilă* du roumain).

assignables au IVe ou au IIIe millénaire av. J.-C.[44]. Mais peut-on alors déjà parler de « Thraces » ?

Les peuples qu'on a quelques raisons de pouvoir appeler « thraces » (à partir du début du IIe millénaire av. J.-C.) ignorent l'écriture jusqu'à la fin de l'époque archaïque. On voit alors apparaître l'alphabet grec (§ 2.1.2) et plus précisément attique, pour les Odryses de l'époque classique. Mais l'écriture resta, semble-t-il, longtemps assez confidentielle et les quelques documents thraces ou supposés tels demeurent, on l'a vu, impénétrables, n'apportant ainsi aucune lumière sur la langue.

Les exégètes du thrace s'appuient donc essentiellement sur les gloses (§ 2.2.1) et sur les noms propres présents

a) dans les inscriptions grecques (depuis le Ve siècle av. J.-C.), puis latines,

b) dans les littératures grecque et latine (§ 2.2.2).

On imagine aisément la difficulté des problèmes soulevés par une telle approche.

Problèmes chronologiques d'abord, pour les matériaux d'origine littéraire : les données, dans ce secteur, vont d'Homère à la fin de l'Antiquité ; l'on ne peut naturellement se fier à la date présumée de l'ouvrage et l'on doit s'interroger sur celle des sources. Quand, par exemple, telle glose d'Hésychius a-t-elle été enregistrée ? ou, tel auteur a-t-il recueilli lui-même l'information ? sinon, à qui en est-il redevable ?

Problèmes formels, surtout : les sources sont le plus souvent non pas thracophones, mais hellénophones[45]. Cela suppose

a) une adaptation morphologique au grec (puis au latin) avec apparition de suffixes grecs, puis latins (voir Beševliev, 38 *sqq.*) ;

b) une intégration phonologique au grec, en fonction des deux systèmes en présence et selon des modalités aujourd'hui bien connues ;

c) des phénomènes de « polarisation » : assimilation d'un radical indigène à un radical grec ou latin phonétiquement proche, processus banal quand deux langues se rencontrent, cf. les anthroponymes thraces en -τοκος assimilés à ceux en -δοκος du grec[46], ou latin *Mucianus* vraisemblablement senti en Thrace comme procédant du radical épichorique Μοκα-/Μουκα- (Beševliev, 38-40).

Certains graveurs d'inscriptions étaient évidemment bilingues ; mais

a) ils ont dû être très tôt prisonniers d'une certaine tradition graphique ;

44. Un sceau d'argile à Karanovo (N.-E. de Plovdiv), une tablette d'argile à Gradešnica (N. de la Bulgarie ; photo chez Neroznak, après p. 112) et trois autres à Tărtăria (Transylvanie), voir Georgiev, *Premier Symposium*, 22, et H. Siegert, *I Traci* (trad. ital. de *Wo einst Appolo lebte*), Milan, 1986, 36-38.

45. Généralement, les formes « latines » semblent être passées (à l'exception, parfois, du suffixe) par un intermédiaire grec.

46. C'est là, à coup sûr, la meilleure explication du flottement -τοκος/-δοκος, si -τοκος est la forme « authentique ». Voir, *infra* § 3.7.1., le point de vue des « mutationnistes ».

b) ne l'auraient-ils pas été qu'ils se seraient de toute façon heurtés à l'inadaptation partielle de l'alphabet;

c) enfin, en raison de l'autonomie du code écrit, une forme graphique peut vivre sa propre vie : si un nom thrace a été enregistré avec un Φ grec à l'époque où ce dernier valait */ph/*, quand celui-ci s'est spirantisé Φ a pu être lu [*f*] (d'où latins PH et F), sans que cela corresponde à une évolution thrace[47].

Pour ajouter à ces difficultés, il faut naturellement songer

a) à l'évolution que la ou les langue(s) indigène(s) n'ont pu manquer de connaître au cours du millénaire sur lequel s'étale notre documentation;

b) à la diversification dialectale qu'ont dû engendrer le morcellement géographique et politique et les vicissitudes de l'histoire.

3.2. Questions de méthode

Les thracologues évoluent donc sur un terrain excessivement mouvant, qu'ils essaient de pénétrer essentiellement par la méthode étymologique. Ils sont souvent conscients du danger et multiplient les mises en garde, refusant fréquemment *a priori* les étymologies «radicales» (*Wurzeletymologien*) fondées sur des ressemblances formelles, pour ne retenir que celles où l'on a réussi à établir le sens du mot par «la méthode combinatoire, sans tenir compte de son étymologie» (Georgiev, 1966, 120).

3.2.1.

Ainsi, pour un théonyme, on propose une interprétation à partir des attributs ou des propriétés de la divinité : Σαβαδιος/Σαβαζιος nom du Dionysos thrace, est censé être l'équivalent sémantique du latin *Liber* (cf. vieux bulgare *svobodĭ* «libre» < **swobhodhios*; Georgiev, 1984, 209). L'épithète thrace d'Asklépios Ζυμδρηνος, Ζυμεδρηνος, Ζυμζδρηνος, Ζυμλυζδριηνος, etc.[48], référerait au «serpent d'eau», principal attribut du dieu (d'où prototype **gh°myu-udr-āno-*, Georgiev, 1983, 1155).

On essaie de procéder de la même façon pour toponymes ou hydronymes.

a. Αξιος est le nom d'un affluent du Danube dans la Dobroudja,

b. Αξιο-οπα/Αξι-ουπολις celui d'une ville aux bouches du Danube,

c. Αξιός/Αξειός celui de l'Axios/Vardar en Macédoine;

a et **b** portent aujourd'hui en bulgare le nom de *Cernavoda* «Eau-noire». L'affluent principal de l'Axios/Vardar s'appelle aujourd'hui *Crna (reka)*

47. Autres cas (celui du **ū*) illustrant éventuellement le même phénomène, *infra* 3.4.2.

48. Seize variantes chez Detschew, 1957, 195-196.

« Rivière-noire ». Αξιος (**a**)/᾿Αξιός - ᾿Αξειός (**c**) doit donc renvoyer à **n̥-ks(e)j* « noir, obscur », cf. le nom de la mer Noire Πόντος Ἄξεινος, d'où Εὔξεινος (Georgiev, 1966, 121). Γερμας, Γερμαη, etc., cité de la région de Pautalia (Ouest de la Bulgarie, près de Kjustendil), porte à présent le nom de *Saparevska Banja* « Bains de Saparevo », qui est connue pour ses sources d'eau chaude : la forme remonterait au même étymon que le grec θερμός « tiède, chaud » (*ibid.*, 122).

3.2.2.

Les noms de personnes, simples ou composés, représentent, on l'a dit, une part importante du matériel. Certains sont extrêmement fréquents : C. Poghirc (*Thracians and Mycenaeans*, 298) relève 360 exemples de Βιθυς, 132 de Τηρης, 115 de Σευθης, 83 de Κοτυς, 70 d'Αυλουζενης, 62 de Δολης, 50 de Διζας, etc.[49]. Une analyse interne, par confrontation des simples et des composés ou des composés entre eux permet d'isoler les différents radicaux. Ainsi, les couples Αυλουκενθος - Αυλουζενης ou Μουκαπορις - Μουκαζενις mettent en évidence les radicaux Αυλου-, -κενθος, Μουκα-, -πορις et -ζενης/-ζενις, les uns figurant toujours en première position (*e.g.* Αυλου-, Μουκα-), les autres en seconde seulement (-ζενης/-ις, -κενθος). Puis, usant raisonnablement de la comparaison, au prix d'une hypothèse minimale, on met le composé thrace en rapport par exemple avec un composé grec supposé génétiquement identique : Δια-ζενις avec Διο-γένης, Διζα-ζενις avec Θεογένης ; ou avec un composé typologiquement identique : Βρια-ζενις[50] avec ᾿Αστυ-γένης, Δεσα-κενθος/Διζα-κενθης avec Θεό-τεκνος. Et de proche en proche on parvient ainsi à fournir une étymologie plausible pour nombre des unités reconnues : Δια- < **diw(o-)*, Διζα-/Δεσα- = grec Θεο- (quelle que soit la forme de l'étymon indo-européen), -ζενης/-ζενις < **genH*$_1$- (grec -γενης), -κενθος/-κενθης < **ken-* (« né récemment », etc.), cf. latin *recens* (Pokorny, *IEW*, 563 sq., **ken-*/3)[51]. Nous empruntons ces exemples à un article, excellent à bien des égards, de Georgiev, *Pulpudeva* 2, 7-19.

On parvient de cette façon à d'indéniables résultats. Bien des thracologues ont simplement le tort de vouloir tout expliquer : est-on certain que le rapport de -πουις/*-puis*/*-pus* (Μουκαπουις, *Mucapuis, Mucapus*) avec le grec παῖς dépasse la simple assonance ? que -πορις/-πουρις ait quelque chose à voir avec latin *puer* ? que Βιθυ- procède de **bwitu* (grec φῖτυ) ? que -ζελμις (cf. Αυλου-ζελμις) doive être rattaché à la racine **kel-* « protéger » (Pokorny, *o.c.*, 553, **kel-*/4) ? Encore ne donnons-nous

---

49. Ces chiffres comptent l'ensemble des variantes.

50. Sur le sens de βρια, voir *infra* § 3.3.

51. Georgiev, 1984, 209.

ici que des exemples où l'écart entre le point de départ supposé et le point d'arrivée n'est pas considérable. Nous avons vu plus haut des étymologies plus «acrobatiques», à propos de Σαβαζιος et de Ζυμδρηνος (§ 3.2.1).

C'est qu'on n'a pas toujours le soutien d'une comparaison plausible, que les modèles sémantiques des anthroponymes sont variés et que, même là où, comme pour les hydronymes ou les toponymes, ils le sont moins, l'étymologiste travaille souvent «sans filet». Que vaut, par exemple, l'interprétation de Μαμουζηνη (nom d'une divinité) par **ma-mnti-ānā* «*Mater montana*»[52], ou de Μυγδονία par **mūs-ghdhōm* «Le Pays des Mysiens»[53] ou **mūko-ghdhōm-* «Le pays marécageux»[54]. Il n'est donc pas étonnant que le chemin des thracologues, qui surestiment généralement la méthode étymologique[55], soit jonché de cadavres. L.A. Gindin (*Dritter Kongress* I, 34-35) souligne que, pour les anthroponymes, sur 482 étymologies proposées par Tomaschek 10 étaient actuellement acceptées, 75 sur 654 pour les toponymes et 34 sur 95 pour les hydronymes. Et les meilleurs, hélas, se laissent prendre à ce jeu, à force de voir reprendre les mêmes équations, cf. C. Poghirc, 1971, 175, qui considère comme assurée la relation de thrace *mezenai* (de lecture d'ailleurs incertaine, *supra* § 2.1.2.b) avec messapien **Menzanas* (= Jupiter) et roumain *mînz* «poulain».

C'est pourtant sur des bases aussi chancelantes qu'on essaie d'établir la grammaire de la ou des langues concernées et, au-delà, la géographie linguistique de la région.

3.3. La thèse du thrace et du daco-mésien

Jusqu'à Detschew, pour les idiomes parlés dans l'aire roumano-bulgare, on parle simplement de «thrace». Russu utilise le terme de thraco-dace, sans envisager une diversité autre que dialectale.

C'est Georgiev qui introduit en 1957[56] l'idée de la bipartition linguistique de la région en thrace proprement dit au Sud du flanc septentrional de

52. Georgiev, 1984, 211.

53. Georgiev, 1966, 176.

54. I. Duridanov, cité par G. Mihailov, 1986, 380-381.

55. Souvent, comme on le verra plus loin, au prix d'une démarche circulaire : l'hypothèse suscite l'étymologie, qui vient ensuite conforter l'hypothèse.

56. A. Vraciu, *Thraco-dacica*, 323, cite sans références le nom d'un précurseur roumain, B.P. Hasdeu, qu'apparemment Georgiev n'aurait pas connu. Si Vl. Georgiev est fréquemment cité, c'est non seulement à cause de l'abondance de sa production, mais aussi parce qu'il est très représentatif de l'«école balkanique» (largement tributaire de l'«école viennoise») et qu'il est très souvent suivi dans ses conclusions (en Bulgarie et ailleurs).

l'Haimos (les Balkans) et daco-mésien ou daco-gète au Nord du Danube, avec (introduite plus tard) une zone intermédiaire entre Haimos et Danube, où l'on aurait parlé «*eine thrakisch-dakische Mischsprache*» (1983, 1191). Les deux langues seraient irréductibles l'une à l'autre.

Les résultats fournis par les étymologies (*infra* § 3.4) lui paraissent confirmés par la répartition de certains toponymes :

– Les noms en -παρα (neutres; -παρος, -παρον, -βαρι, -πωρον)[57], en -βρια (féminins)[58] et en -διζα (neutres)/-διζος[59] n'apparaissent jamais au Nord du Danube, alors que ceux en -δαυα/-δαβα/-δεβα (toujours neutres ?)[60] ne sont, à une exception près, présents qu'en Dacie et dans les deux Mésies.

D'autres ont essayé d'approfondir et de justifier cette thèse, *e.g.* K. Vlahov à partir de l'anthroponymie[61] ou C. Poghirc qui, au prix d'une enquête incluant noms de lieux, de montagnes, de cours d'eau, de tribus, de dieux, de plantes, conclut que sur 3 017 entrées prises en compte, 36 seulement procèdent d'un radical panthrace (Dacie + Mésies + Thrace)[62].

On ne peut certes nier l'intérêt des répartitions ainsi entrevues. Mais :

– Il faut souligner l'énorme disproportion qui existe entre le matériel fourni par la Thrace proprement dite et celui qu'ont donné la Dacie et les deux Mésies : Poghirc (*l.c.*) opère, par exemple, avec plus de 1 100 anthroponymes pour la Thrace, 195 pour la Mésie Inférieure, entre 50 et 60 pour la Mésie Supérieure et environ 90 pour la Dacie.

---

57. Une cinquantaine du type Αθυπαρα, Βενδιπαρα, Κρασαλοπαρα (un cas en *-phara* : *Breirophara*). Le radical est d'abord rattaché (1966) à un thème **barā* avec le sens de «rivière, fleuve» et rapproché du grec βόρβορος «fange, bourbier» (en réalité, sans doute une formation expressive !); puis Georgiev hésite entre «Market», «Fluss» et «Dorf» (1983); voir *e.g.* Georgiev, 1966, 139 et 176-177; 1983, 1196.

58. Une quinzaine d'exemples du type Μεσαμβρια, Πολτυμβρια ou Σαλαμβρια. Βρια est donné comme l'équivalent de πόλις par Strabon (7.6.1) et Etienne de Byzance (*s.v.*, Μεσημβρια) et comme valant ἡ ἐπ' ἀγροῖς κώμη par Hésychius (*s.v.* βριαν). Le mot est rattaché à un radical **wrijā* (tokh. A *ri*, B *rije* «ville, forteresse»); voir Georgiev, 1966, 139 et 178; 1983, 1196; Velkova, 48-50.

59. Une dizaine d'exemples tels qu'*Orudiza*, Κιστιδιζος, Τυροδιζα : étymon probable **dh(e)igh-*, cf. grec τεῖχος, voir Georgiev, 1966, 178; 1983, 1196.

60. Une cinquantaine de cas comme Βρεγεδαβα, Καρσιδαυα, Ιταδεβα, cf. Hésychius λεβα· πόλις ὑπὸ Θρακῶν, rectifié en δεβα par Tomaschek (Velkova, 56-60) et généralement rattaché à **dheH*$_1$*-w-* ou **dhH*$_1$*-w-* «place»; voir Georgiev, 1966, 173-175; 1983, 1211. Sur la répartition des quatre types, voir les cartes de Georgiev, 1966, 136-137 et 172. G. Mihailov, *Pulpudeva* 5, 8-10, s'interroge sur le sens primitif de ces radicaux compte tenu de l'urbanisation tardive de la région. Sw. Janakiewa, *Dritter Kongress* I, 250-253, croit pouvoir montrer que ceux en -βρια sont les plus anciens.

61. *Ethnogenèse*, 193-197; ainsi les noms en Αυλου- (rarement Αυλο-), premier élément de plus de 130 composés, se rencontrent sur le cours moyen de l'Hébros et sur la mer Noire (pays des Odryses et des Besses), mais non à l'Ouest ni dans le bassin du Nestos; les quelques cas observés en Mésie et Dacie auraient été apportés par l'immigration.

62. 1976, 335-347.

– Si la Dacie connaît des noms aussi caractéristiques que Βυρεβιστας Δεκεβαλος ou *Scorilo*, près de la moitié des anthroponymes relevés dans cette région se retrouvent en Thrace ; mais on a tendance à sous-estimer le fait en raison de la mobilité géographique des noms de personnes : le principe est juste en soi, encore que le départ entre ce qui est propre au pays et ce qui ne l'est pas comporte une part de subjectivité.

– En revanche, ne surestime-t-on pas la toponymie ? Les noms en -διζα n'apparaissent, par exemple, qu'à l'Est et au Sud-est de la Thrace et personne ne songe à faire de cette zone une région linguistiquement autonome (Papazoglou, 78-79).

– Les limites toponymiques ne sont pas aussi nettes que le voudrait Georgiev : dans la région de Sofia-Pirot-Kjustendil, une aire anciennement thrace, cinq des six noms en -δαβα/-δεβα/-δεβαι ne sont connus que par Procope (VIe siècle ap. J.-C.) et pourraient être attribués à la venue (IIIe siècle ap. J.-C.) de Daces de la Dacie transdanubienne ; mais le sixième, *Desudaba*, livré par Tite-Live, ne saurait s'expliquer ainsi ; or il est dans le pays des Μαῖδοι, tribu dont personne ne conteste le caractère thrace (Papazoglou, *l.c.*). En outre, comment expliquer la présence de *Pulpudeva* (= Φιλιππόπολις/Plovdiv) au cœur de la Thrace proprement dite [63] ? Enfin, outre que l'élément *-dava* se retrouve entre Balkans et Danube, on observe en pays triballe et mésien des hydronymes à « phonétisme thrace » [64] (Αθρυς, Αρτανης, *Utus, Timachus, infra* §§ 3.7 et 3.7.1), ce qui contraint Georgiev à imaginer la zone intermédiaire évoquée ci-dessus.

Bref, Vl. Georgiev soulève un problème réel, mais qui doit être traité avec plus de circonspection. La division qu'il prétend ainsi fonder est-elle confirmée par les données linguistiques les plus fiables ?

3.4. Quelques traits phonétiques très discutés[65]

Les graphies sont caractérisées par une grande instabilité et leur interprétation dépend largement des étymons proposés.

63. Autre anomalie : selon Mihailov, *Pulpudeva*, 5, 9, *Pulpudeva*, si l'on en juge par un nom médiéval de Plovdiv (*Plovdin*ъ), aurait connu une seconde forme, *Pulpudina* ; or l'élément *-dina* (généralement féminin ?) figure dans toute une série de toponymes « sur les deux rives du Bas Danube, du côté pontique de la région gète » (C. Poghirc, 1971, 172), cf. *Amlaidina*, Ασβολοδεινα, Βασσιδινα, etc. (*id.*, 1976, 338).

64. Selon les critères de Georgiev, naturellement.

65. Pour une vue d'ensemble sur les résultats phonétiques proposés, voir (avec beaucoup d'oscillations, d'approximations et de repentirs) Russu, 93-100 et 151-152 [162-172] ; Georgiev, 1966, 129 (thrace), 143 (daco-mésien) et 146-148 (thrace et daco-mésien) ; *id., Dritter Kongress* I, 209-212 (thrace) ; *id.*, 1983, 1163-1174 (thrace), 1181-1184 (daco-mésien), 1189-1191 (thrace et daco-mésien) ; Katičić, 143-151 (qui suit dans l'ensemble la thèse de Georgiev).

3.4.1.

**o* : Russu, 94-95 [164-165], croit naturellement à un aboutissement unique de **o*, qui aurait conservé son timbre, avec flottement *o/a* tardif et dialectal. Georgiev prétend que **o* est passé à *a* en thrace (1983, 1165) comme en daco-mésien (1182) et il allègue des formes comme Αυλου-ζανος (où -ζανος viendrait de **gonos*) ou Δια-ζενις (= grec Διο-γένης). Σπαρτακος/*Spartacus* serait, face au fréquent Σπαρτοκος/-δοκος, la forme authentiquement thrace. Le flottement *o/a*, rare et peu sûr en thrace, trahirait une évolution tardive en daco-mésien de [*a*] (< **a*, **o*) en [*o*] [66].

3.4.2.

**u* et **ū* conserveraient leur timbre en thrace et en daco-mésien selon Georgiev (1983, 1165-1166 et 1182), cf. *Utus*/Ουτως (affluent du Danube) de **udōr* (grec ὕδωρ) ou βρουτος/βρυτος (sorte de bière) de **bhrūtos*. Mais **ū* serait tardivement passé à [*y*], puis [*i*], cf. pour une localité du S.-E. de la Thrace, *Surascele, Syracella, Syrascel(l)e, Sirogellis* (Detschew, 1957, 486), et, pour le daco-mésien, le nom des Mésiens Μυσοί > Μοισοί [67]. Or, si les étymologies proposées sont exactes, **ū* est rendu graphiquement par OY (latin U) ou Y (latin Y). Quand ce dernier graphème a correspondu à [*y*] dans le système graphique grec, en vertu de l'autonomie de l'écrit (*supra* § 3.1) il a pu tout simplement être lu [*y*], d'où OI recouvrant alors la même articulation; puis Y et OI en étant arrivés à valoir [*i*], ils étaient susceptibles d'êtres remplacés par I ou tout autre graphème de même valeur. Les flottements de l'orthographe risquent donc de ne pas correspondre à une évolution thrace ou daco-mésienne.

3.4.3.

Réalisation du point d'appui vocalique accompagnant les sonantes dans certains contextes : Russu, 95 [166], hésite à se prononcer, faute d'étymologies sûres. Pour Georgiev, en thrace (1983, 1167-1168) °*l* et °*r* auraient abouti à *ul (ol)* et *ur (or)*, tandis que °*n* donnait *un* ou *in* : cf. Δορζας, *Durze*, etc. (nom de personne) rapporté à **dh°rs-* (grec θρασύς) ou βρυνχος, cithare thrace d'après Hésychius, censé remonter à **bhr°mkos* (cf. grec φόρμιγξ). En daco-mésien, *°*r* > *ri* et *°*n* > *a*, cf. Κρισος (rivière) < **kr°so-* «noir», ou Αξιος (*supra* § 3.2.1). Les résultats dépendent donc toujours largement des étymologies avancées et nous ne citons ici que quelques-unes des moins aventureuses.

---

66. Cette mutation serait conditionnée par la proximité d'une labiale ou d'un liquide, cf. en Dacie *Potaissa* pour *Patavissa*, voir C. Poghirc, 1989, 296-306, qui retrouve le même écart notamment dans les hydronymes daces qui survivent en roumain (cf. Αλουτας > *Olt*).

67. I.v. Bredow, *Dritter Kongress* I, 249, évoque là pour le thrace une mutation dialectale.

3.5.

Les labiovélaires et les aspirées constituent l'un des rares secteurs où il y a consensus, malgré des divergences sur le détail et les incertitudes entourant bien des étymologies proposées : les premières (*k^w et *g^w/*g^wh) se sont délabialisées (d'où *k, g*) [68] ; les secondes (**bh, *dh, *gh*) ont perdu leur appendice « soufflé » (d'où *b, d, g*), et ceci dans le thraco-dace de Russu (96-97 [167-168]) ou dans les deux langues de Georgiev (1983, 1169-1171, 1173, 1183), cf. thrace γεντον « viande » < *g^w*hento-* (glose, Velkova, 55-56), toponyme dace Γερμι-ζερα/*Germisara* et variantes < *g^w*herm-* (Detschew 1957, 103), glose thrace βρουτος/βρυτος < **bhrūto-* (Velkova, 53-55), etc.

En thrace, le produit de la délabialisation des labiovélaires serait soumis à la mutation consonantique (*infra* § 3.7) : ainsi le toponyme Κελλαι/Κελλη remonterait à *g^w*elnā* « source » (avec *g^w > *g* > *k*), selon Georgiev, 1983, 1173. Pour le dace, à partir de Γερμι-ζερα > Ζερμι-ζεργα (*supra*) ou de Ταπαι (les actuelles « Portes de Fer » sur le Danube) qui remonterait à **tok*^{w}*oy* « Strömungen », le même Georgiev (*ibid.* 1183) prétend que le produit *k, g* des labiovélaires s'est « später » palatalisé devant *e/i*, mais est devenu dialectalement *p, b* devant *a, o, u* : il est bien évident que, si ce traitement « dialectal » s'avérait exact, il ne pourrait être intervenu « später » (c'est-à-dire après intermédiaire *k, g*), mais devrait procéder directement de *k^w, *g^w, *g^wh.

3.6. Le thrace et la dichotomie « *centum-satem* »

Nous croyons que cette dichotomie repose non sur l'existence d'une série de dorsales palatales qui auraient été palatalisées ici et se seraient confondues avec les dorsales vélaires là, mais sur la palatalisation conditionnée de l'unique série de dorsales dans certains dialectes indo-européens [69] : c'est sans doute l'une des premières variations dialectales indo-européennes.

Cela dit, malgré les flottements graphiques, il semble bien que le thraco-dace ou le thrace et le daco-mésien appartiennent à la catégorie

68. Une survivance sporadique de la labiovélaire est parfois évoquée, cf. nom de plante dace κοαδαμα, rapporté à **(a)k*^{w}*ā* « eau », voir Detschew, 1957, 551, et Katičić, 144.

69. Voir, à ce sujet, l'excellent développement de Georgiev, 1966, 21-48, qui, après avoir montré la surévaluation traditionnelle du traitement *satem* (voir indépendamment, Cl. Brixhe, *Le lingue indoeuropee di frammentaria attestazione*, Pise 1983, 131), conclut à la nécessité de bannir la division des langues indo-européennes en *satem* et *centum*.

des langues indo-européennes qui connaissent cette palatalisation (groupe *satem* selon la terminologie traditionnelle) : les occlusives concernées y apparaissent sous les graphies Σ, Ζ, Δ [70], cf. *e.g.* -ζενης = grec -γενής, Ρησος (nom d'un roi mythique, puis nom de personne) mis en rapport avec lat. *rēx* etc., -διζα (élément toponymique) de **dhigh-* (grec τεῖχος), ζετραια = grec χύτρα selon Hésychius et Pollux (Velkova, 67-68) de **gheutr-* (cf. grec χέω, χύτρα), etc. ; en tout cas, il paraît y avoir unanimité sur ce point (Katičić, 143-144).

3.7. La mutation consonantique (*Lautverschiebung*)

Contre Russu (96 [166-167]), qui croit au maintien intact des occlusives indo-européennes voisées (*b, d, g*) et non voisées (*p, t, k*), Georgiev suppose que, s'il en est bien ainsi en daco-mésien (1983, 1183), le thrace connaît, lui, une mutation consonantique : **b, *d, *g > p, t, k, *p, *t, *k > ph, th, kh*, cf. hydronymes *Utus* < **udōr* ou Αθρυς < **ētrus* « rapide » (*ibid.* 1170-1172). Pour la série sourde, l'aspiration, qui pourrait avoir été faible (d'où les oscillations graphiques), ne serait pas intervenue après *s* [71], cf. σκαλμη « épée, couteau thrace » selon Sophocle et Hésychius, qui serait l'avatar de **skolmā* (Velkova, 81-82). Comme nous l'avons vu plus haut (§ 3.5), l'aboutissement des labiovélaires serait, lui aussi, soumis à la mutation.

3.7.1.

C'est là une thèse qui vient de loin : esquissée par Mladenov en 1915, elle a été développée par Detschew et Georgiev [72]. Il convient de noter :

a. Que la règle présenterait de nombreuses exceptions, tant dans les gloses (cf. γεντον, βρουτος/Βρυτος déjà cités) que dans l'onomastique : l'élément anthroponymique -κενθος/-κενθης, qui remonterait à **kent-* (§ 3.2.2) comporte toujours un κ ; les noms en -πορις (ou variantes), qui aurait le même étymon que lat. *puer*, présentent toujours un π ; -δοκος/-δακος/-τοκος, autre élément anthroponymique, qui viendrait de **dokos* (mais voir *supra* § 3.1), offre parfois la forme -θακος (inattendue dans le cadre de cette théorie ; Detschew, 1957, 145) ; entre Balkans et Danube, région daco-mésienne, qui devrait donc avoir échappé à la mutation, on rencontre des formes comme *Utus*/Ουτως (cf. *supra*), ce qui conduit Georgiev à imaginer une zone tampon, d'abord thrace, puis recouverte par les Daco-mésiens (*supra* § 3.3).

70. Georgiev, 1983, 1172-1173 (thrace) et 1183 (daco-mésien).
71. Georgiev, 1984, 211.
72. Historique de la question chez Duridanov, *LB* 31 (1988) 1-2, 57-64.

En vérité, cette hypothèse est, avec l'évolution supposée de **o* en **a* et les migrations tribales ou ethniques, l'un des outils les plus évidents de la manipulation linguistique et historique : ainsi, à cause sans doute de la permanence de *k* dans -κενθος/-κενθης, Georgiev, qui s'en était toujours tenu jusque-là à l'étymon **kent*-, imagine (1984, 211) un prototype **(s)kentor*- (!). Selon le même, en raison de son *o*, -τοκος/-δοκος (pratiquement inconnu en Dacie et en Mésie) serait «(proto)phrygien», (*Pulpudeva* 2, 7). À cause de la persistance de la sonore, seraient macédoniens Αργιλος, nom d'une ville de Bisaltie (à l'Ouest du cours inférieur du Strymon; < **arg*-) et Αναδραιμος, site thrace sur lequel fut fondée Amphipolis, traduit en grec par Ἐννέα Ὁδοί «les Neuf Routes» et qui pourrait venir de **enwa drom*- (grec δρόμος)[73]. On comprend donc le scepticisme de certains, ainsi Mihailov, 1986, *l.c.*, et 1987, 19, n. 33.

Incontestablement, il faut reprendre le dossier, vérifier les formes, préciser leur chronologie, les replacer dans leur contexte historique, puis, sur la base d'une bonne philologie, se poser quelques questions : que vaut la tradition latine, le plus souvent tributaire de la grecque? Que valent aux IIe-IIIe siècles de notre ère Φ, Θ et X? Quel peut être le poids du grec à un moment ou à un autre de son histoire phonologique et graphémique, étant donné que le matériel est généralement fourni par un discours grec? etc. C'est seulement après avoir tenté de résoudre ces problèmes qu'il conviendra d'essayer d'élaborer une théorie rendant compte de l'ensemble des graphies, à partir peut-être de deux hypothèses simples, qui, excluant la mutation en question, peuvent se combiner :

a. absence d'aspirées en thrace (cf. Φιλιππόπολις traduit par *Pulpudeva*, avec même sort pour *ph* et *p* du grec), provoquant chez les Thraces l'assimilation de l'aspirée grecque à leur propre sourde et l'interchangeabilité des signes correspondants (T/Θ, K/X, Π/Φ)[74];

b. les sonores thraces ont pu avoir été des sonores fortes et les sourdes des fortes avec une aspiration comparable à celle qu'on observe encore dans l'allemand contemporain en diverses positions[75].

---

73. Georgiev, cité par Mihailov 1986, 382-383, qui souligne à juste titre que ces toponymes sont anciens et que les Macédoniens n'interviennent qu'assez tard dans la région (Ve siècle). Le caractère non thrace de ces formes est encore affirmé par Duridanov, *LB* 31 (1988) 1-2, 60.

74. Le même phénomène se rencontre partout où le grec rencontre une langue sans aspirées, en Asie Mineure (Cl. Brixhe, *Essai sur le grec anatolien au début de notre ère*[2], Nancy 1987, 110-113 et 157), en Egypte (Brixhe, *REG* 103, 1990, «Bulletin de dialectologie grecque», 223, n° 46), chez le Scythe des *Thesmophories* d'Aristophane (Brixhe, in *L'Etranger dans le monde grec*, Nancy, 1988, 119-123).

75. Cf. M. Philipp, *Phonologie de l'allemand*, Paris, 1970, 56-57.

3.8. Le thraco-dace ou le thrace et le daco-mésien parmi les langues indo-européennes

Étant donné la fragilité des diverses constructions proposées quant à la situation linguistique ancienne de l'aire roumano-bulgare, vaut-il la peine de s'interroger sur la position dialectale de la ou des langue(s) concernée(s) ?

Disons simplement que l'hypothèse d'une unité thraco-phrygienne, inaugurée par Kretschmer en 1896, semble aujourd'hui généralement abandonnée, même si le phrygien est souvent présent dans les spéculations des thracologues. Le thrace et le daco-mésien ou le thraco-dace, langue(s) «*satem*», et le phrygien, qui ne l'est pas, sont le plus souvent considérés comme deux ou trois langues différentes, qui, si elles ont des affinités, sont irréductibles à un même groupe linguistique antérieur. Mais il est actuellement difficile d'accorder quelque crédit à l'idée (fort répandue) [76] selon laquelle, dans l'Antiquité (tardive), elles auraient constitué avec l'illyrien et le macédonien un «*Sprachbund*» soumis à la pression du grec.

3.9. La survie de la ou des langue(s)

Russu (110, 153 [193]) note que les derniers anthroponymes thraces apparaissent au VIe siècle ap. J.-C. ; si l'on en croit deux indications antiques, selon lesquelles il semblerait que le dialecte besse était encore parlé à la même époque dans un monastère à proximité de la mer Morte et dans un autre dans le Sinaï (Russu, 110, 153 [194-195]), on peut imaginer que la langue du peuple qu'Hérodote (V 3) considérait comme «le plus nombreux de la terre après les Indiens» a survécu au moins marginalement jusque-là. Ensuite, nous n'en entendons plus parler.

76. Cf. Georgiev, 1984, 214, qui évoque, en outre la proche parenté du thrace et du pélasgique (substrat prégrec, objet de maintes spéculations aventureuses) et l'existence de nombreux points communs au thrace, au daco-mésien (qui serait l'ancêtre de l'albanais) et aux langues balto-slaves.

Bibliographie

Ancient Bulgaria : Ancient Bulgaria. Papers Presented to the International Symposium on the Ancient History and Archaeology of Bulgaria (1983). – (Univ. of Nottingham 1981), ed. by A.G. Poulter, I, Nottingham.

Ancient Macedonia : Ancient Macedonia (1986). – IV (Papers Read at the Fourth International Symposium Held in Thessaloniki, September 1983), Salonique.

BEŠEVLIEV V. (1970). – *Untersuchungen über die Personennamen bei den Thrakern*, Amsterdam.

BONFANTE G. (1955). – «A Note on Samothracian Language», *Hesperia* 24, 101-109.

Contributions : Contributions au IVe Congrès international de thracologie. (1984). – (Rotterdam 1984), Sofia.

DANOV Chr. M. (1976). – *Altthrakien*, trad. par G. Minkova, Berlin – New York.

DETSCHEW D. (1952). – *Charakteristik der thrakischen Sprache* (en bulgare et en allemand), Sofia; 2e édit. *LB* II (1960), 145-213.

– (1957). – *Die thrakischen Sprachreste*, Vienne; 2. Auflage mit Bibliographie 1955-1974, von Ž. Velkova (Österr. Akad. Wiss., Schriften der Balkankommission. Linguistische Abt. vol. XIV), Vienne, 1976.

Deuxième Symposium : Actes du deuxième Symposium international de thracologie. (1980). – (Rome, 1979), Milan.

Dritter Kongress : Dritter Internationaler Thrakologischer Kongress zu Ehren W. Tomascheks. – (1984). (Vienne 1980), Sofia.

Ethnogenèse : L'ethnogenèse des peuples balkaniques (1971). – (Symposium International sur l'ethnogenèse des peuples balkaniques, Plovdiv, 1969), Sofia.

FRASER P.M. (1960). – *Samothrace* 2.I. *The Inscriptions on Stone*, Londres.

FRIEDRICH Joh. (1932). – *Kleinasiatische Sprachdenkmäler*, Berlin.

GEORGIEV Vl. (1957). – *Trakijskijat ezik (La langue thrace)*, Sofia, 1957; 2e édit. revue et augmentée *Trakite i tehnijat ezik (Les Thraces et leur langue)*, en bulgare avec résumé en français, Sofia, 1977.

– (1966). – *Introduzione alla storia delle lingue indeuropee*, Rome, 1966 (édition utilisée ici), repris par *Introduction of the History of the Indo-european Languages*, Sofia 1981.

– (1983). – *ANRW* II.29.2, Berlin – New York 1983, «Thrakisch und Dakisch», 1148-1194, «Thrakische und dakische Namenkunde», 1195-1213.

– (1984). – «Die thrakische Sprache im System der indoeuropäischen Sprachen», *Dritter Kongress* I, 207-215.

HODDINOTT R.F. (1990). – *Les Thraces* (trad. franç. par Cl. Sorel), Paris.

KATIČIĆ R. (1976). – *Ancient Languages of the Balkans* (= Trends in Linguistics, State-of-the-Art Reports 4-5), La Haye-Paris.

KRETSCHMER P. (1896). – *Einleitung in die Geschichte der griechischen Sprache*, Göttingen, 172-243.

LEHMANN K. (1955). – «Documents of the Samothracian Language», *Hesperia* 24, 93-100.

– (1960). – *Samothrace* 2.II *The Inscriptions on Ceramics and Minor Objects*, Londres.

MIHAILOV G. (1984). – «L'épigraphie thrace», *Dritter Kongress* II, 208-212.

– (1986). – «L'onomastique de l'aire thraco-macédonienne», *Ancient Macedonia*, 377-392.

– (1987). – «Les inscriptions dans le trésor de Rogozen», *LB* 30. 1, 5-19.

NEROZNAK V.P. (1978). – *Paleobalkanskije jazyki*, Moscou.

PAPAZOGLOU F. (1978). – *The Central Balkan Tribes in Pre-Roman Times (Triballi, Autariatae, Dardanians, Scordisci and Moesians)*, Amsterdam.

POGHIRC C. (1971). – «Réflexions sur les problèmes du daco-moesien», *Ethnogenèse*, 171-177.

– (1976). – «Thrace et daco-mésien : langues ou dialectes ?», *Thraco-dacica*, 335-347.

– (1989). – «Considérations chrono-géographiques sur l'oscillation a/o en thrace et en daco-mésien», *Thracians and Mycenaeans*, 296-306.

Premier Symposium : Actes du Premier Symposium international de thracologie (1978). – (Rome 1977), Milan.

Pulpudeva : Semaines philippopolitaines de l'histoire et de la culture thrace 1 (Plovdiv, 1974), Sofia, 1976 ; 2 (Plovdiv, 1976), Sofia, 1978 ; 3 (Plovdiv 1978), Sofia, 1980 ; 4 (Plovdiv, 1980), Sofia, 1983 ; 5 (Plovdiv, 1982), Sofia, 1986.

Quatrième Congrès : Contributions au IVe Congrès international de thracologie (1984). – (Rotterdam, 1984), Sofia.

RUSSU I.I. (1969). – *Limba traco-dacilor* (en roumain avec résumés en russe et en français), Bucarest, 1959, seconde édition Bucarest 1967 ; repris, augmenté ou modifié sous le titre *Die Sprache der Thrako-Daker*, Bucarest ; ouvrage cité d'après l'édition de 1959, avec, entre crochets droits, références au texte de 1969.

Thracians and Mycenaeans (1989). – (Proceedings of the Fourth International Congress of Thracology, Rotterdam 1984), ed. by J.G.P. Best and N.M.W. de Vries, Leyde – New York – Copenhague – Cologne.

Thraco-dacica (1976). – Recueil d'études à l'occasion du IIe Congrès International de thracologie (Bucarest 1976), édit. par C. Preda, A. Vulpe et C. Poghirc, Bucarest.

TOMASCHEK W. (1980). – *Die alten Thraker* I *in* Sb. Akad. Wien 128, IV. Abh. (1893), 1-130 ; II, *in* Sb. Akad. Wien 130, II Abh. (1893), 1-70 ; III *in* Sb.

Akad. Wien 131, I. Abh. (1894), 1-103. Ces trois études ont été reprises sous le titre *Die alten Thraker. Eine ethnologische Untersuchung*, Vienne.

Troisième Symposium : Actes du IIIe Symposium international de thracologie (1982). – (Palma de Mallorca 1981), Rome.

VELKOVA Ž. (1986). – *The Thracian Glosses. Contribution to the Study of the Thracian Vocabulary*, Amsterdam.

CHAPITRE 10

LE MACÉDONIEN

Claude BRIXHE, Anna PANAYOTOU

1.

Historiquement, ceux à qui l'on donne le nom de Macédoniens, correspondent à un ensemble de tribus unifiées, selon la légende, par les Téménides au VII[e] siècle av. J.-C., non sans que subsistent jusqu'au temps de Philippe II de petits royaumes macédoniens, semi-dépendants.

On admet généralement que ces tribus étaient descendues du massif du Pinde, à l'Ouest, et avaient occupé progressivement vallées et plaines, leur siège restant longtemps le bassin de l'Haliacmon, où était précisément située leur ancienne capitale, Aigéai.

À partir de là, elles n'ont cessé de gagner du terrain, au Nord-Ouest et au Nord aux dépens des Illyriens, des Péoniens, des Phrygiens, au Sud aux dépens des Pières, à l'Est, surtout, aux dépens de diverses tribus thraces et de colonies grecques, le plus souvent ionophones (voir *Le phrygien*, §§ 0 et 0.1 ; *Le thrace*, § 1.2.1).

Avec Alexandre I[er] (*ca* 498-454 av. J.-C.), les Macédoniens atteignent le Strymon. C'est, selon l'opinion commune, Archélaos (413-399 av. J.-C.)

qui déplace la capitale d'Aigéai à Pella, sanctionnant ainsi le déplacement du centre de gravité du royaume et ses ambitions orientales. Philippe II donne un nouvel élan à cette expansion : le Nestos est atteint [1].

Ce fleuve constituera d'ailleurs la frontière orientale de la province romaine de Macédoine en 148/147, qui s'étend ainsi du Pinde à l'Ouest au Nestos à l'Est, de la Thessalie et de l'Egée du Sud à un peu au-delà de Stoboi au Nord [2].

Les Macédoniens occupent donc un territoire de plus en plus vaste, incluant des Illyriens, des Thraces, des Phrygiens, des Péoniens. Mais, eux-mêmes, qui étaient-ils et quelle langue parlaient-ils ?

2. LES THÈSES MODERNES

L'histoire moderne a partagé la Macédoine entre trois pays : la Grèce, la Yougoslavie et la Bulgarie, et, de la seconde moitié du XIX[e] siècle à nos jours, l'ethnicité des Macédoniens et l'origine de leur langue ont parfois servi d'argument en faveur de prétendus droits ancestraux sur la région. L'appréciation des faits est donc trop souvent affectée par des interférences politiques nationalistes.

2.1.

L'essentiel de la documentation étant linguistique, le problème de la langue est naturellement au cœur du débat. Mais le dossier documentaire est, somme toute, fort mince et apparemment très hétérogène. Cette minceur et cette hétérogénéité ont favorisé la prolifération de thèses divergentes. Avec prudence, on peut être tenté de regrouper les réponses données en quatre types [3], susceptibles de se recouper partiellement, avec diverses variantes qu'il est hors de question de développer ici.

---

1. Les Macédoniens vont même naturellement au-delà, jusque sur la Propontide et le Pont. Mais le lit du Nestos resta, à l'Est, la limite territoriale la plus stable.

2. Pour l'histoire et la géographie de la région, voir N.G.L. Hammond, *A History of Macedonia* I. *Historical Geography and Prehistory*; II. *550-336 B.C.* (en collaboration avec G.T. Griffith); III. *336-167 B.C.* (en collaboration avec F.W. Walbank), Oxford 1972, 1979, 1988; et le livre magistral de F. Papazoglou, *Les villes de Macédoine à l'époque romaine* (*BCH*, Suppl. XVI), Athènes-Paris 1988 (avec les testimonia sur la période préromaine également).

3. Voir Panayotou 1990, 129-134 (§§ 5 à 5.6.7) et 370-379 (§ 8.47.4), avec bibliographie.

En fonction, parfois, des préjugés précédemment évoqués, ou selon qu'est pris en compte l'ensemble du matériel dans sa diversité ou qu'on essaie de le ramener à l'unité ou encore qu'on veut atteindre «une couche originellement macédonienne» (Masson, 178) le macédonien est considéré comme :

a. Une langue mixte à base illyrienne : thèse inaugurée en 1825 par K.O. Müller[4] et défendue récemment encore par G. Bonfante[5].

b. Une langue mixte constituée à partir d'un dialecte grec, avec influences illyriennes et thraces; c'est, pour l'essentiel, l'hypothèse de Kretschmer[6] ou de Schwyzer[7], par exemple.

c. Un dialecte grec, avec diverses appréciations quant à sa place parmi les parlers helléniques (cf. O. Hoffmann et sa thèse éolienne)[8], avec éventuellement évocation de substrats non grecs, cf. M. Sakellariou, qui croit déceler des isoglosses avec les dialectes du Nord-Ouest, le groupe éolien et l'arcadien, et des traits substratiques ressortissant au thrace et au pélasgique[9].

d. Une langue autre que le grec, parfois sentie comme proche de lui, parfois considérée comme appartenant au même groupe préhistorique que le grec et le «thraco-phrygien» : les Macédoniens seraient un peuple non grec historiquement hellénisé; c'est la thèse d'I.I. Russu[10], par exemple; c'est celle vers laquelle semblent incliner, avec beaucoup de réserves des linguistes comme A. Meillet[11] ou O. Masson (178-179)[12].

La bonne synthèse de R. Katičić (108-116) reflète parfaitement ces hésitations : soulignant l'absence d'une solution définitive, il se demande (115-116) si les Macédoniens – comme opposés aux peuples subjugués – ne parlaient pas un dialecte grec (thèse **c**), sans exclure l'hypothèse d'une langue étroitement apparentée au grec (thèse **d**).

---

4. *Über die Wohnsitze, die Abstammung und die ältere Geschichte des makedonischen Volkes. Eine ethnographische Untersuchung*, Berlin.

5. *RAL* (8ᵉ série) 42.5-6 (1987), 83-85.

6. *Einleitung in die Geschichte der griechischen Sprache*, Göttingen, 1896, 283-288.

7. *Gr. Gr.* I, 69-71.

8. *Die Makedonen, ihre Sprache und ihr Volkstum*, Göttingen, 1906.

9. In *Macedonia, 4000 Years of Greek History and Civilization*, Athènes, 1983, 44-63.

10. *Ephemeris Dacoromana* 8 (1938), 105-232 (synthèse p. 158).

11. *Aperçu d'une histoire de la langue grecque*[7], Paris, 1965 (avec bibliographie mise à jour et complétée par O. Masson), 61.

12. Telle était la position d'O. Masson en 1967/1968. Il se range à présent sous la thèse précédente et considère désormais le macédonien comme un dialecte grec (*per litteras*).

3. Les témoignages antiques[13]

L'hellénisme des Macédoniens a été contesté dès l'Antiquité : mais on sait avec quelle facilité les Grecs pouvaient utiliser l'épithète « barbare » (à propos de la langue, des mœurs, etc.)[14].

Les témoignages apparemment les plus précis nous viennent d'Hérodote :

1) Μακεδνόν et Δωρικόν seraient deux appellations successives du même *ethnos*, respectivement quand il était encore dans le massif du Pinde et quand il arriva dans le Péloponnèse (I 56 ; cf. VIII 43) ; les Macédoniens auraient donc constitué un rameau des Doriens.

2) La dynastie royale des Téménides serait venue d'Argos (VIII 137-140).

3) Contre les Grecs qui voulaient l'écarter des jeux Olympiques, Alexandre Ier fait reconnaître ses origines grecques : ἐκρίθη τε εἶναι Ἕλλην (V 22).

Les passages antiques consacrés à la langue peuvent tous s'expliquer par une hypothèse grecque, sauf peut-être Quinte-Curce, *Histoire d'Alexandre*, VI 11.4, qui laisse entendre que le macédonien avait besoin d'un interprète pour être compris[15] ; mais cette remarque est en contradiction avec un autre développement du même auteur (VI 9. 34-36), où les propos d'Alexandre laissent croire que le macédonien pouvaït – fût-ce avec difficulté – être compris de tous les Grecs[16]

L'information la plus précise semble être fournie par Tite-Live (31.29.15), où l'on parle des *Aetolos, Acarnanas, Macedonas, ejusdem linguae homines* (v. *infra* § 5.1.2).

Avec le temps, μακεδονίζειν ou μακεδονικός, concernant la langue, ont pu tout simplement référer à la *koiné* courante (par rapport au standard attique), à l'expansion de laquelle les Macédoniens étaient associés politiquement et culturellement[17].

---

13. Ils sont rassemblés par Panayotou, 1990, 97-98 (§ 2.1, pour l'ethnicité) et 121-128 (§ 4, pour la langue) ; voir Katičić, 104-108.

14. Voir, Cl. Brixhe, *Lalies* 9 (1990), 27.

15. En fait, cette information ne saurait être retenue comme un argument définitif contre l'héllénisme des Macédoniens ; entre deux Grecs (un Athénien et un Pamphylien, par exemple) y avait-il nécessairement intercompréhension ?

16. Panayotou, 1990, 123-124.

17. A. Panayotou, dans *Ἡ γλῶσσα τῆς Μακε δονίας*, Athènes, 1992, 192-194.

4. Le matériel

La Macédoine est restée longtemps un pays marginal par rapport au monde grec, économiquement arriéré, où la vie pastorale était sans doute largement répandue dans les zones montagneuses. L'alphabétisation devait, au mieux, être limitée à l'aristocratie et à son entourage immédiat. Pour l'époque archaïque, on n'attend donc *a priori* aucun document privé écrit.

Il est vrai qu'à la même époque, dans le reste de la Grèce, l'écriture est surtout l'apanage de la cité, qui s'en sert pour écrire le droit; or nous sommes ici en monarchie et dans un système aristocratique; on n'a donc pas à s'étonner de l'absence d'épigraphie publique.

Ainsi, toutes les conditions étaient réunies pour que la province n'émerge épigraphiquement qu'assez tard.

Les premiers documents écrits laissés par les Macédoniens sont des légendes monétaires (à partir de la fin du VIe siècle av. J.-C.), peu éclairantes pour notre sujet.

Quand les Macédoniens commencent à écrire (fin du Ve – début du IVe siècle av. J.-C.), ceux qui s'expriment le font en attique (*koiné*). La ou les langue(s) indigène(s) est (sont) occultée(s), parce que restreinte(s) à la communication privée ou... déjà défunte(s).

Mais il ne faut pas désespérer de voir cette nuit documentaire trouée par l'éclair de quelque découverte épigraphique.

N'a-t-on pas trouvé récemment à Pella une *defixio* assignable au IVe siècle av. J.-C. et écrite en un dialecte grec original [18] ? Certes, un document de ce type n'est pas, on le sait, nécessairement rédigé dans la langue locale; certes le rédacteur (la rédactrice; ou la bénéficiaire) n'était pas nécessairement macédonien (la population de Pella n'était pas homogène) et l'on doit assurément attendre une seconde trouvaille linguistique identique pour se prononcer définitivement [19]. Mais l'originalité du parler, avec des isoglosses doriennes et nord-occidentales – ce qui n'étonne pas, compte tenu de la situation géographique de la Macédoine – rend probable son authenticité macédonienne.

---

18. Emm. Voutyras *Ἑλληνικὴ διαλεκτολογία* 3(1992-1993), 43-48 (version préliminaire d'un long article à paraître).

19. Ce sera peut-être possible avec la découverte (postérieure à la rédaction de ce chapitre) de plusieurs documents non rédigés en koiné, à Aiané, sur le cours moyen de l'Haliacmon : une épitaphe peinte, un graffite sur vase, un compte, une inscription gravée sur le bord d'un pithos, datables du Ve siècle av. J.C. ; un graffite de la première moitié du IVe siècle av. J.C.

4.1.

Hormis peut-être ce document et ceux qu'annonce la note 19, on n'a donc aucun texte écrit en macédonien ; mais des gloses nous sont parvenues sous cet ethnique, presque essentiellement par Hésychius.

Leur nombre varie beaucoup d'un auteur à l'autre [20] : on y inclut parfois des unités sans attribution ethnique, parce qu'elles viennent d'Amérias, lexicographe d'origine macédonienne ou parce qu'en raison de leur consonantisme elles semblent macédoniennes.

Parmi les gloses livrées avec l'étiquette « macédoniennes », certaines peuvent concerner des mots grecs de large diffusion répandues par les Macédoniens ou qui avaient pris un sens particulier en Macédoine ou plus simplement dans la *koiné* courante (« vulgaire ») : ainsi κοράσιον (= παιδίσκη, κόριον, etc.) est donné tantôt comme macédonien (*Schol. B ad Il.* XX *404*), tantôt comme « vulgaire » (« ne se disant pas », Phrynichos), tantôt comme étranger (Photius, *Lex., s.v.* παιδισκάριον) ; selon Phrynichos encore, ῥύμη était employé à Athènes avec le sens d'ὁρμή, mais avec celui de στενωπός chez οἱ νῦν ἀμαθεῖς, puis il ajoute δοκεῖ δέ μοι καὶ τοῦτο μακεδονικὸν εἶναι. C'est que, comme on l'a vu *supra* § 3, il y a eu avec le temps évolution de la notion de « macédonien ». L'appréciation portée sur un lexème peut donc être tributaire de sa date d'enregistrement par les lexicographes.

Il n'est pas sans intérêt de noter que, sur les 95 gloses retenues pour la présente étude, plus de la moitié appartiennent à une communication relativement restreinte, à des secteurs du lexique où, dans une aire linguistique de quelque étendue (et à plus forte raison, dans le domaine grec, morcelé géographiquement et politiquement), on enregistre fréquemment des écarts d'une région à l'autre : épithètes religieuses (14), faune (12), termes de la vie rurale (noms de plantes, particularités géographiques,... 16), lexique culinaire (5), vocabulaire technique militaire (7).

Aucun renseignement ne nous est fourni sur la grammaire, sinon :

a) l'affirmation, par Apollonios Dyskolos et Eustathe, de l'existence (non vérifiée épigraphiquement) de nominatifs masculins singuliers -α en macédonien, et

b) peut-être ὔμμες (= ὑμεῖς) que, selon Athénée, VII 323b, Strattis mettait dans la bouche d'un ξένος (un Macédonien ?) dans sa comédie *Pausanias* ou *Les Macédoniens*.

4.2.

Les inscriptions en *koiné* et les auteurs anciens, grecs ou latins, complètent notre documentation en nous fournissant des anthroponymes, des toponymes, des épithètes divines, des noms de mois.

20. Ainsi 153 chez J.N. Kalléris, 1954 ; environ 140 chez Schwyzer, *Gr. Gr.* I, 69.

Les mêmes inscriptions en *koiné* présentent, en outre, certaines particularités phonétiques ou morphologiques (§ 5), qui sont trop rarement prises en compte pour l'appréciation de la situation linguistique et de sa genèse.

Le matériel en notre possession nous renseigne donc avant tout sur le lexique. Mais, de la grammaire – essentielle pour situer une langue – nous n'entrevoyons que quelques bribes : quelques traits phonétiques, quelques éléments morphologiques (surtout dans le domaine de la dérivation).

5. LE MACÉDONIEN

Si nous avions eu la place nécessaire pour développer les différentes hypothèses avancées concernant le macédonien, on aurait perçu une extrême confusion quant à l'usage de ce terme. C'est pourquoi, avant toute investigation, il n'est peut-être pas inutile de poser, au moins théoriquement, les principes de l'emploi le plus juste du mot.

La Macédoine est, on l'a vu, un territoire primitivement hétérogène :

1) S'agissant de la population, il faut se garder du fantasme de l'origine, se garder de vouloir remonter le temps en droite ligne pour arriver aux « Macédoniens », aux seuls à mériter ce nom. L'entité humaine macédonienne est vraisemblablement composite, faite d'osmoses successives.

2) Si derrière le matériel disponible, on découvre une langue **A**, avec substrat (adstrat ?) **B**, quelle est celle qui devra être qualifiée de « macédonienne » ? *A priori*, la langue **A**, susceptible d'avoir été le parler de ceux qui ont unifié la province, politiquement et, sans doute dans une large mesure, linguistiquement aussi. **A** risque d'avoir été la langue de la couche ou de l'ethnie hégémonique, qui, sans qu'on puisse en évaluer l'expansion primitive, a fini par s'imposer, « retouchée » par **B**, comme la langue de l'entité humaine composite évoquée plus haut. De l'examen des apports de **B**, on pourra éventuellement tirer des indications : – sur les rapports socio-politiques primitifs entre porteurs de **A** et porteurs de **B** ; – sur la possible survie de **B**.

5.1. Les éléments grecs

Le matériel décrit *supra* § 4 livre des traits incontestablement grecs (§ 5.1.1), dont certains autorisent des rapprochements avec un dialecte documenté (§ 5.1.2).

5.1.1. Traits grecs sans coloration dialectale

Lexique

On y observe :

– Des mots grecs qui se rencontrent ailleurs, avec à peu près le même sens ou un sens voisin, cf. le vocatif ἄππα, terme enfantin « macédonien » (selon l'*Etym. Magnum, s.v.* ἄττα) à l'adresse d'un plus âgé, qui appartient à un type (celui de *Lallnamen*) sans frontière, banal dans ce registre ; ainsi, en grec même ἄππας, attesté épigraphiquement avec le sens de « religious official » et donné comme valant τροφεύς par Hésychius (voir le lexique *L.S.J.*).

– De banals régionalismes [21] dont on n'est même pas certain qu'ils ne se soient pas développés dans le cadre de la *koiné*, cf. le κοράσιον cité plus haut ou encore ἄργος = τὸ πεδίον, macédonien, thessalien et παρὰ τοῖς νεωτέροις d'après Strabon (VIII 372).

– Certains mots qui s'écartent de la forme grecque standard par un trait phonétique mineur, cf. ἰνδέα · μεσημβρία, Μακεδόνες (Hésychius), où ἰνδέα correspond au féminin d'ἔνδιος [22].

– Des lexèmes non attestés ailleurs, mais clairement grecs : ainsi τελεσιάς danse guerrière (Athénée, XIV 629d ; cf. τελέω) ; ἀόρτης = ἄγγος δερμάτειον ἱματίων chez les Macédoniens (Hésychius), cf. les banals ἀορτή « aorte » mais aussi « havresac/besace » et ἀορτήρ « baudrier » ; Μιμαλλόνες dit des femmes au comportement viril, cf. μιμηλός, μιμηλάζω ; ainsi peut-être également δάρυλλος·ἡ δρῦς ὑπὸ Μακεδόνων (Hésychius) : un dérivé de δρῦς avec *a* anaptyctique ? Nous sommes là dans un domaine où, dans toutes les langues, les appellations varient souvent d'une région à l'autre.

Anthroponymie

L'immense majorité des anthroponymes attestés en Macédoine est grecque : Ἀλέξανδρος, Φίλιππος, Πτολεμαῖος, Ἀντιγόνη, Λάανδρος, Λανίκα, Λαοδίκα, Λᾶγος, etc. (Masson, 178).

Toponymes et ethniques

– Il n'est pas absolument certain, mais seulement probable que le nom même des Macédoniens soit d'origine grecque : Μακεδών et variantes,

21. Le lexique d'Hésychius fourmille de tels régionalismes, cf. *e.g.* ἄδρυα· πλοῖα μονόξυλα, Κύπριοι. Λέγονται δὲ καὶ οἱ ἐν τῷ ἀρότρῳ στῦλοι. Σικελοὶ δὲ ἄδρυα λέγουσι τὰ μῆλα ; βάβακοι = τέττιγες en Elide mais βάτραχοι dans le Pont ; ἄμπελος = αἰγιαλός à Cyrène, etc.

22. Voir en dernier lieu Panayotou, 1990, 192 (§ 8.2.5.).

Μακεδνόν (ἔθνος, Hérodote, I 56, VIII 43), Μακέτης (v. *L.S.J. s.v.*), cf. homér. μακεδνός « long, élevé » [23].

– La plupart des toponymes connus sont interprétables par le grec (cf. Αἰγεαί, Πέλλα, Αἰανή...) [24], en particulier ceux du bassin de l'Haliacmon. En tout cas, il ne semble pas y avoir en Macédoine plus de noms de lieux prégrecs qu'en une autre région grecque.

Cette constatation est importante, car

1) ce sont les « maîtres » qui « nomment », et

2) la seule présence d'une dynastie royale grecque régnant sur un peuple étranger (hypothèse maintes fois évoquées) [25] pourrait difficilement avoir eu cet effet.

Epithètes divines

Avec Θαύλιος ἢ Θαῦλος· Ἄρης Μακεδόνιος (Hésychius), ou Θούριδες· νύμφαι, μοῦσαι, Μακεδόνες (Hésychius), par exemple, nous sommes encore en pays grec.

Calendrier

Dans le calendrier macédonien, les mois Δῖος, Ἀπελλαῖος (notons l'archaïsme du vocalisme *e*), Ἀρτεμίσιος, Λῷος, Δαίσιος doivent leur nom au grec.

5.1.2. Traits à caractère dialectal

La koiné écrite en Macédoine (et parfois les gloses) présente une coloration particulière, qui ne peut s'expliquer que si l'on suppose qu'elle a recouvert une autre variété de grec.

Ce dialecte avait conservé au */ū-ŭ/ son timbre primitif, cf. Κουναγίδᾳ, épithète d'Héraclès dans une dédicace de Styberra en Derriopos (191 ap. J.-C., *SEG* II 436) [26].

Le */ā/ ne s'y était pas fermé en /ē/ et là où la forme attendue, celle de la koiné attique, présente normalement un η, elle offre fréquemment un α en Macédoine [27] : dans le lexique (cf. l'épigraphique χειριστᾷ [28]), mais

23. D'ailleurs une origine non grecque du/des mot(s) ne saurait prouver l'origine non grecque de la population : nombre de peuples/de pays doivent leur nom à une langue étrangère, cf. « France/français ».

24. Cf. Vl. Georgiev, *Introduzione alla storia delle lingue indeuropee*, Rome, 1966, 189-190 ; Katičić, 114.

25. Cf. Kalléris, 1976, 506, n.1.

26. Cf. S. Düll, *Die Götterkulte Nordmakedoniens in römischer Zeit*, Munich, 1977, 342, n° 135.

27. Voir Panayotou, 1986, 420-426 ; 1990, 210-213 (§ 8.7), 407-409 (§ 9.1.1.11) et 426-427 (§ 9.1.2.12) ; Brixhe-Panayotou, 250-251 et 254.

28. Amphipolis, fin du IIIe/début du IIe siècle av. J.-C., L. Moretti, *ISE*, vol. II, 108-114, n° 114, col. A³3.

surtout dans l'onomastique (Εὐρυδίκα, ᾿Αντιγόνα, Φιλώτας, ᾿Ικκότας, ᾿Αμύντας, etc.); dans les finales (cf. les exemples précédents), mais aussi dans les radicaux, cf. l'appellatif ἄγημα, corps d'élite (grec non ionien-attique ἀγέομαι)[29], noms de personnes ῾Αδίστη, ῾Αδέα, ῞Αδυμος, etc.[30].

Quand deux langues différentes, même apparentées, se rencontrent, la langue dominante ne souffre pas semblable altération. Ainsi le phrygien, langue proche du grec, a conservé intacts */ū-ŭ/ et */ā/ indo-européen (v. *Le phrygien*, § 2.1.1); il survit au moins jusqu'au Ve siècle de notre ère; or, en dehors de formes comme ᾿Αμύντας ou Φιλώτας, primitivement véhiculées précisément par des Macédoniens qui en ont imposé la finale, la *koiné* de Phrygie est toujours conforme, sur les deux points ici évoqués, au modèle attique. Κουναγίδας et ῾Αδίστη ne peuvent guère s'expliquer que par la superposition, à un moment donné, de deux variétés de grec, de deux formes identiques à un détail près, non aléatoire, mais entrant dans un jeu de correspondances régulières : *koiné* /ε̄/[31] (H) ~ dialecte /ā/ (A), *koiné* /ȳ-y̆/[32] (Y) ~ dialecte /ū-ŭ/ (OY).

Cette situation est précisément celle qu'on aperçoit partout où la *koiné* recouvre un dialecte grec qui a conservé un /ā/.

A pour H reste fréquent jusqu'assez tard après le début de notre ère. Le trait est alors, à une ou deux exceptions près, limité à l'onomastique; il n'est probablement plus lié à une réalité dialectale vivante, mais est simplement, sur bases traditionnelles, volonté d'affirmer une identité.

Nous avons donc très certainement là une preuve que, au moment de l'atticisation, au moins une large partie de la population parlait une variété locale de grec. Le pays a donc connu une période de diglossie sans doute assez prolongée, avec dialecte réservé à la communication privée et n'accédant que rarement à l'écrit.

D'autres traits identifiés dans la *koiné* des inscriptions, dans les gloses et dans le texte inédit de Pella devraient nous permettre d'affiner notre perception de cet adstrat (puis substrat) grec, très certainement de type occidental : en raison de la position géographique de la Macédoine, on ne s'étonnera pas de rencontrer des isoglosses avec les parlers du Nord-Ouest (document de Pella) et avec le thessalien, ainsi le traitement de **-sm-* dans ὔμμες (*supra* § 4.1), si authentique, ou encore éventuellement la fermeture sporadique de /ō/ en /ū/, non conditionnée en thessalien, mais apparemment liée à la proximité d'une nasale en Macédoine, cf. l'anthroponyme Κάνουν = Κάνων ou ἀκρουνοί· ὅροι ὑπὸ Μακεδόνων (Hésychius)[33].

---

29. Cf. P. Chantraine, *BSL* 61 (1966), 161.

30. Voir Chantraine, *o.c.*, 164-165; Panayotou, 1986, 425.

31. Puis /ē/, /ī/, et enfin /i/.

32. Puis /ī-ĭ/ et enfin /i/.

33. Où ἀκρουνοί devrait correspondre à un *ἀκρωνοί; sur la question, voir A. Panayotou, dans *Poikila* (Meletemata 10), Athènes, 1990, 211-212.

Quelle que soit sa source, Tite-Live n'était donc peut-être pas loin de la vérité quand il disait qu'Etoliens, Acarnaniens et Macédoniens parlaient la même langue [34].

5.2. Eléments non grecs

5.2.1.

L'essentiel du matériel connu est donc grec. Mais on y trouve des éléments hétérogènes, que seules des contorsions linguistiques permettraient de ramener au grec. Aucun d'entre eux ne concerne la grammaire. Étant donné la position sociale probable, par rapport au grec (dominant), de la ou des langues [dominées(s)] qui les ont fournis, le contraire eût été étonnant.

Sont concernés :

– Des anthroponymes : Σαβαττάρας [35], Ἀρραβαῖος/Ἐρρεβαῖος, Ἀρριδαῖος/Ἐρριδαῖος [36] Ἀμάδωκος [37], etc.

– Des toponymes, *e.g.* Βισαλτία, Ὀδομαντική, Σιντική...

– Des épithètes ou noms divins : Δάρρων· Μακεδονικὸς δαίμων ᾧ ὑπὲρ τῶν νοσούντων εὔχονται (Hésychius, confirmé épigraphiquement par une dédicace de Pella), Σαυᾶδαι·Σαῦδοι, nom macédonien des Silènes selon Amérias (Hésychius), ou Ζειρηνίς· Ἀφροδίτη ἐν Μακεδονίᾳ (Hésychius).

– Des noms de mois : Αὐδ(υ/ου/ο)ναῖος, Γορπιαῖος, Ξανδικός.

– Des appellatifs : ἄλιζα· ἡ λεύκη τῶν δένδρων (Hésychius) [38], ἀβροῦτες · ὀφρῦς (Hésychius), δάνον = Θάνατον (Plutarque, *Moralia*, 22c).

Si parfois on en entrevoit l'origine, par exemple thrace pour Ἀμάδωκος ou Σαυᾶδαι, nous sommes le plus souvent dans l'obscurité : en raison de l'histoire, on doit supposer l'intervention de plusieurs langues.

La comparaison éclaire parfois l'origine indo-européenne du mot, sans permettre de préciser l'ethnie qui l'a fourni, cf. γόδα· ἔντερα (Hésychius) en face de sanskrit *gudám*, *gudás* «intestin». Mais que vaut

34. Malgré Katičić (108), qui, à cause de sa date, considère le témoignage comme suspect.

35. V. O. Masson, *RPh* 53 (1979), 244.

36. V. O. Masson, in Φιλίας Χάριν (Mélanges E. Manni) IV, 1980, 1480-1481 ; A. Panayotou, *o.c.* (n. 33), 208-209.

37. Epigramme funéraire d'Aigéai (350-340 a.C.), *SEG* 35, 789.

38. Plutôt «la lèpre blanche des arbres» que (avec correction ἡ λεύκη τὸ δένδρον) «le peuplier blanc»)? Cf. Katičić, 110, n. 161.

un rapprochement avec indo-européen **alisā* « aulne » (Pokorny, *IEW*, 302, *s.v. el-/1*), si ἄλιζα désigne la lèpre blanche des arbres ?

Beaucoup d'incertitudes entourent une bonne partie des formes.

5.2.2.

Cependant, mérite un examen particulier un petit groupe de mots qui avaient déjà attiré l'attention des Anciens et ont fait couler beaucoup d'encre depuis le siècle dernier : ils présentent le signe d'une occlusive sonore là où le grec a celui d'une occlusive aspirée[39].

Il s'agit :

– D'appellatifs : ἀβροῦτες· ὀφρῦς (Hésychnius), ἀδραία· αἰθρία (Hésychius), δάνον = θάνατον (Plutarque), δῶραξ· σπλήν (Hésychius), κεβαλήν/κεβλήν = κεφαλήν (*Etym. Magnum, s.v.* Βέροια ; etc.).

– D'une épithète divine : Δάρρων (*supra* § 5.2), cf. grec θαρσέω/θαρρέω, τὸ θάρσος/θάρρος.

– D'un nom de mois : Ξανδικός (grec Ξανθικός).

– D'un toponyme : Βέροια = Φέροια selon l'*Etym. Magnum.*

– Et surtout d'anthroponymes : Βερενίκα/Βερνίκα, Βερεννώ (grec φέρω), Κέββας, Κέβων (grec κεφαλή), Βίλος, Βίλιστος, Βίλιππος (grec φίλος), etc.

On a donné du phénomène diverses interprétations[40]. Une seule, cependant, nous semble recevable : la langue responsable de cet apport s'est séparée du grec quand les deux idiomes avaient encore la série indo-européenne des sonores aspirées, **bh, *dh, *gh* ; celles-ci y ont par la suite perdu leur appendice « soufflé » (d'où *b, d, g*)[41], tandis qu'en grec elles perdraient leur voix[42].

5.2.2.1.

Pour certains des éléments du dossier, la présence d'une aspirée dans l'étymon indo-européen est assurée : ἀβροῦτες[43], ἀδραία, Βερενίκα, Δάρρων... Elle l'est aussi pour Βίλος etc., si l'on en croît l'équivalent grec Φίλος et malgré l'absence de correspondants sûrs dans les autres langues i.-e.

D'autres n'ont pas d'étymologie : δῶραξ, Ξανδικός...

Enfin le nom de la « tête », κεβαλή/κεβλή (et anthroponymes correspondants), a tout à fait l'allure d'un hybride : puisqu'il remonte à **ghebh(e)l-*, on attendrait *γεβαλή/γεβλή. C'est d'ailleurs peut-être la

39. Dossier chez Panayotou, 1990, 357-378 (§ 8.47).

40. Résumées par Panayotou, *ibid.*, 370-378 (§ 8.47.4).

41. C'est assurément le traitement le plus banal, dans les langues indo-européennes.

42. C'est, par exemple, la thèse de Schwyzer, *Gr., Gr.*, I, 70.

43. Où l'élargissement apical peut n'être pas indo-européen (Masson, 177), mais être seulement l'un des aspects de l'intégration morphologique du mot au grec.

forme attendue que nous avons avec γαβαλάν· ἐγκέφαλον ἢ κεφαλήν (Hésychius, sans ethnique). Plutôt que d'expliquer *kebl-* comme procédant, par exemple, de **gebl-* avec dissimilation, on peut se demander si, à partir de l'équation « grec Φ, Θ, X = macédonien B, Δ, Γ », on n'a pas :

– vu surgir des étymologies populaires, cas de Βέροια ?

– vu fabriquer, par les Macédoniens eux-mêmes (et les grammariens ?), de faux macédonismes, cf. les faux dialectismes qui apparaissent dans le monde grec lors du déclin des dialectes et qui étaient destinés à marquer l'identité de leurs auteurs [44].

5.3.3.

La plupart des formes retenues précédemment sont attestées ou risquent de remonter à une époque où le grec de Macédoine (dialecte, puis dialecte et *koiné* attique) avait conservé leur occlusion aux aspirées et aux sonores : l'écart entre Φ et B, quand on passe de Φερενίκη à Βερενίκα/Βερνίκα, reflète alors la distance entre [*ph*] et [*b*] et ne saurait s'interpréter comme un voisement [45], /*ph*/ n'ayant pas de contrepartie voisée. Mais à basse époque, quand sonores et aspirées sont devenues des spirantes, les deux séries constituent une corrélation de voix : *v* (graphie B), δ (Δ), γ (Γ) ~ f(Φ), θ (Θ), x (X) ; les échanges entre B et Φ ont alors une tout autre signification : marquant le passage de [*v*] à [*f*] ou *vice versa* (de la sonore à la sourde ou *vice versa*), ils entrent naturellement dans un ensemble de flottements graphiques entre signe de la sonore et signe de la sourde et sont à placer aux côtés de K-Γ (*k-g*, Δικαία-Διγαία), Π-B (*p-b*, ὑπό-ὑβό) [46], etc.

À partir du matériel épigraphique tardif, risquent donc de s'introduire dans le dossier des échanges parfois apparemment semblables à ceux qui ont été examinés dans le paragraphe précédent, mais qui en réalité n'ont rien à voir avec eux et sont sans signification ethnique.

5.3.4.

C'est une des langues subjuguées qui a communiqué les quelques formes du § 5.2.2 au dialecte grec de Macédoine. Il s'agit là d'un traitement banal, puisque partagé, dans la région, par l'illyrien (apparemment) [47], le thrace [48] et le phrygien [49].

La phonétique (et notamment le vocalisme) des termes concernés semble aussi conservatrice que celle du grec. Cela pourrait rendre suspecte

---

44. Voir, par exemple, R. Hodot, *Lalies* 9 (1990), 59-60.
45. Ainsi M. Hatzopoulos, *BCH* 111 (1987), 406 *sq*.
46. Exemples empruntés à M. Hatzopoulos, *l.c.*
47. Voir Katičić, 170-171.
48. Cf. *Le thrace*, § 3.5.
49. Cf. *Le phrygien*, § 2.1.2.

leur attribution à un dialecte thrace ou illyrien. Devrait-on les assigner au(x) parler(s) des tribus phrygiennes non passées en Asie Mineure ? Ce n'est pas impossible.

On a souligné ailleurs la proximité du grec et du phrygien. Serait-elle à nouveau illustrée par Βίλιστος/Βιλίστα, qui impliquerait *a priori* la même formation de superlatif pour le grec et la langue donneuse ? On ne peut l'exclure ; mais, on ne peut exclure non plus un habillage « macédonien » de Φίλιστος, à partir de l'équation Φ = B (*supra* § 5.2.2.1 *in fine*).

Enfin, la langue pourvoyeuse de Βερενίκα/Βερνίκα est probablement éteinte au moment de l'émergence épigraphique de la province (fin du V[e] siècle av. J.-C.)[50]. Sinon, on comprendrait mal pourquoi il n'y a pas davantage d'échanges Π-Φ, T-Θ et K-X dans les textes en *koiné* ; en effet, quand le locuteur d'une langue sans aspirée (cas du phrygien, par exemple) apprend le grec, il confond la sourde et l'aspirée de ce dernier et les signes correspondants deviennent interchangeables, cf. « Le phrygien », § 2.1.2.

Ce parler ne fut évidemment pas le seul pourvoyeur d'éléments non grecs. Mais notre méconnaissance des langues régionales autres que le grec interdit une discrimination.

6. SIGNIFICATION DES APPORTS NON GRECS AU DIALECTE GREC

Le dialecte grec de Macédoine reflétait vraisemblablement la genèse du peuple macédonien, constitué par osmoses successives avec hégémonie politique et culturelle de tribus hellénophones. Comme on l'attend pour une langue dominante, ce grec ne fut probablement pas affecté dans sa structure ni dans sa grammaire : le trait examiné *supra* § 5.2.2, par exemple, ne faisait qu'augmenter un peu la fréquence des occlusives sonores.

Seuls lexique et onomastique absorbèrent des éléments non grecs. La qualité de cet apport peut projeter quelque lumière sur les relations entre ethnies en présence.

Sont concernés :

a. non seulement la toponymie, ce qui ne surprend pas : même si ce sont les « maîtres » qui « nomment », il reste toujours des résidus antérieurs à eux ;

b. non seulement la vie agricole, le domaine culinaire, ce qui ne surprend pas davantage : les techniques correspondantes peuvent être

50. Ainsi déjà Brixhe-Panayotou, 256

étroitement liées à tel groupe humain et survivre au changement de langue avec incorporation de leur lexique au parler vainqueur;

c. mais aussi le secteur religieux, le calendrier, marque d'une indéniable influence culturelle;

d. et l'onomastique personnelle : presque tous les terroirs grecs présentent des noms que le grec ne peut expliquer. Ce qui est intéressant ici, c'est de voir les anthroponymes non grecs remonter jusqu'au haut de la hiérarchie sociale : on connaît la fortune de Βερενίκη, Ἀρραβαῖος était, entre autres, porté par un roi des Lyncestes, Ἀρριδαῖος/Ἐρριδαῖος était également un nom dynastique (cf. le frère d'Alexandre le Grand).

L'impact d'une langue sur une autre est directement proportionnelle à l'influence politique et culturelle de ses porteurs. La qualité de l'apport linguistique des non-Grecs (cf. surtout les points **c** et **d**) indique qu'ils ont sans doute joué un rôle non négligeable dans la genèse de l'entité historique macédonienne. Illyriens, Thraces et Phrygiens n'ont pas toujours été expulsés, mais, restés en place ou seulement déplacés, ils ont été progressivement absorbés, sans ségrégation, non sans peut-être avoir parfois contribué à la constitution de l'aristocratie du royaume, notamment sans doute à travers des mariages mixtes.

Le grec de Macédoine, dont l'expansion a dû plus ou moins suivre celle de l'Etat, a donc incorporé au patrimoine onomastique grec des noms qui ne l'étaient pas.

Tout naturellement, à un moment donné et face aux Autres (grecs ou non), même quand on fut passé à la *koiné*, ces noms ont été sentis par les Macédoniens comme marque d'une identité, d'où par exemple leur popularité dans la famille royale et l'aristocratie, d'où peut-être aussi la tentation d'en fabriquer de nouveaux à partir d'équations phonétiques reconnues comme spécifiques.

Bibliographie

BRIXHE Cl., PANAYOTOU A. (1988). – «L'atticisation de la Macédoine : l'une des sources de la *koiné*», *Verbum* 11, 245-260.

KALLÉRIS J.N. (1954-1976). – *Les Anciens Macédoniens. Etude linguistique et historique* I-II, Athènes.

KATIČIĆ R. (1976). – *Ancient Languages of the Balkans* (Trends in Linguistics, State-of-the-Art Reports 4), La Haye – Paris.

MASSON O. (1968). – *Annuaire 1967-1968 de l'EPHE, IVe section*, Paris, 176-179.

PANAYOTOU A. (1986). – Γλωσσικὲς παρατηρήσεις σὲ μακεδονικὲς ἐπιγραφές, *Ancient Macedonia* IV (IVe Symposium, Salonique 1983), Salonique, 413-429.

– (1990). – *La Langue des inscriptions grecques de Macédoine (IVe s. a.C. – VIIe s. p.C.). Phonétique, phonologie et morphologie*, thèse de doctorat soutenue devant l'Université de Nancy II en octobre 1990.

CHAPITRE 11

L'ALBANAIS

Shaban DEMIRAJ

1.

L'albanais est une langue indo-européenne parlée aujourd'hui en Albanie (plus de 3 millions de locuteurs), dans l'ex-Yougoslavie (Kosovo-Metohija, Macédoine et Monténégro – environ 2 millions) et en Grèce (au sud de la frontière commune avec l'Albanie), ainsi que par nombre d'émigrés en Turquie, aux États-Unis d'Amérique, en France, en Syrie, en Australie, etc. Il y a aussi de nombreux îlots linguistiques de date ancienne en Italie du Sud, en Sicile et en Grèce (Attique, Péloponnèse, Salamine, etc.), où l'on parle un albanais archaïque fortement influencé là par l'italien, ici par le grec. Le mot *albanais* (en latin *arbanensis/albanensis*) tire son origine de l'ancien ethnonyme/toponyme *Árbën/Arbër* < *Árban* (« Albanie »), dont les habitants s'appelaient *arbënesh/arbëresh* et parlaient l'*arbënisht/ /arbërisht*. Ce dernier mot a été remplacé au seizième siècle par l'adverbe *shqip*, sur lequel sont formés les dérivés *shqip-tar* « albanais », *Shqip-ni/Shqip-ëri* « Albanie », etc.

2.

Tous les parlers albanais se laissent grouper en deux dialectes principaux, le guègue au nord et le tosque au sud du pays. Leurs traits linguistiques communs sont beaucoup plus nombreux que leurs différences dialectales, qui consistent surtout en un nombre limité de phénomènes phonétiques dont les plus importants sont : en guègue, la nasalisation des voyelles devant les consonnes nasales ; en tosque, le rhotacisme du *-n-* intervocalique en *-r-* et le changement de *-a-* et de *-e-* en *-ë-* (semblable au français *-eu-* dans *peur*) devant consonne nasale. Exemples : *zã* ~ *zë* < **zan* « voix », *zan-i* ~ *zër-i* « la voix », *frẽ* ~ *frë* « frein » (< lat. *frēnum*), *fren-i* ~ *frër-i* « le frein », etc. Ces changements phonétiques, qui ne se rencontrent que dans les vieux mots hérités et les emprunts au grec ancien et au latin ont cessé d'être actifs, selon toute vraisemblance, dès le huitième siècle de notre ère. Il est donc improbable que ces changements phonétiques de l'albanais aient quelque connexion historique avec les phénomènes correspondants du roumain, qui sont moins anciens.

Dans le domaine de la grammaire, il faut souligner que l'infinitif périphrastique du type de *me ba* « faire » qui se rencontre dans les parlers guègues fait défaut au dialecte tosque, soit que ce dernier n'ait jamais développé une telle forme verbale, soit qu'il l'ait laissée tomber en désuétude au cours du temps ; en l'état actuel des études, on ne saurait décider entre les deux thèses [1].

3.

En ce qui concerne l'origine de l'albanais, différentes hypothèses ont été avancées jusqu'ici : origine illyrienne, thrace, daco-mysienne, etc. Ceux qui ont présenté des hypothèses sur l'origine non illyrienne de l'albanais ont cherché à les appuyer par des arguments qui, à leur avis, prouveraient que la migration des ancêtres des Albanais vers l'Albanie de l'époque historique s'est faite à partir de la partie orientale ou de la partie centrale des Balkans. Parmi leurs arguments, il faut souligner surtout les concordances lexicales anciennes entre l'albanais et le roumain, ainsi que l'hypothèse

---

1. Voir Shaban Demiraj, *Grammatikë historike e gjuhës shqipe* (Tirana, 1986), p. 964 *sq*.

selon laquelle l'illyrien aurait été une langue du type *centum*, tandis que l'albanais est une langue du type *satəm*. Mais la plupart des spécialistes admettent au contraire que l'illyrien était une langue *satəm*. On ne peut, de toute manière, soutenir le caractère *centum* de l'illyrien en s'appuyant sur des noms propres comme *Acra-banus, Ves-cleves*, etc., où la tectale palatale était suivie de *-r-* ou de *-l-*, car en une telle position une tectale palatale se réalise comme vélaire même en albanais. Quant à la présence de nombreuses correspondances lexicales anciennes entre l'albanais et le roumain, elle peut être expliquée plus rationnellement par le voisinage ancien des ancêtres des Albanais et des Roumains, qui a été interrompu à la suite de l'arrivée des Slaves dans les Balkans (VI-VIIe siècles) [2].

4.

En revanche, l'origine illyrienne de l'albanais peut être étayée par plusieurs arguments, dont les principaux sont les suivants :

1) L'évolution phonétique de toponymes anciens de l'Albanie comme *Durrachion* > *Durrës* (avec accent initial en albanais), *Aulona* > *Vlonë* ~ *Vlorë* (avec rhotacisme en tosque, cf. § 2), etc., a suivi le cours normal de l'évolution historique du système phonétique de l'albanais, ce qui montre que ces toponymes n'ont jamais cessé depuis l'Antiquité d'être en usage chez les ancêtres des Albanais.

2) Il y a en albanais des emprunts au grec ancien (dialecte dorien) et au latin qui datent d'une période antérieure à notre ère, ainsi *mokënë/mokërë* « meule », du grec ancien (dorien) μᾱχανᾱ́.

3) L'histoire n'a enregistré aucune migration massive de l'est à l'ouest des Balkans qui pourrait témoigner d'une migration relativement ancienne des ancêtres des Albanais de l'est des Balkans vers leur pays actuel. Donc, selon toute vraisemblance, l'albanais représente un développement récent de ce qui à date ancienne était l'illyrien du sud. La transition graduelle de l'illyrien à l'albanais a dû avoir lieu au plus tard entre le cinquième et le sixième siècle de notre ère [3].

2. V. Shaban Demiraj, *Gjuha shqipe dhe historia e saj* (Tirana, 1988), p. 44 *sq.*

3. *Ibid.*, p. 103 *sq.*

5.

L'albanais, comme le roumain et le lituanien, est connu par des documents écrits depuis le XVI[e] siècle. Le plus ancien livre en albanais a été publié en 1555. Dès lors cette langue a été cultivée en ses deux dialectes. Mais depuis 1972 on a pris la décision d'employer partout une langue littéraire unifiée, qui a été adoptée même par les Albanais de l'ex-Yougoslavie.

6.

L'albanais se range parmi les langues indo-européennes les plus évoluées. Mais malgré cette évolution frappante il a conservé de nombreux traits linguistiques hérités d'un passé ancien. Parmi ces traits il faut relever :

1) l'apophonie ;
2) le système des trois genres ;
3) la flexion synthétique (nominale et verbale) ;
4) la présence de nombreux mots d'origine indo-européenne, etc.

À l'inverse, parmi les innovations il faut relever :

1) dans le système phonétique : la fixation de l'accent, l'élimination de l'ancienne opposition entre voyelles brèves et voyelles longues, l'évolution des tectales et des consonnes aspirées indo-européennes ;

2) dans la morphonologie : la métaphonie et certains autres phénomènes analogues ;

3) dans le système morphologique : la création d'un thème de pluriel dans les substantifs, le système des articles, la déclinaison au début du mot, les formes verbales analytiques, les formes impersonnelles, etc. ;

4) dans le domaine de la syntaxe : la reprise des compléments d'objet, l'ordre des mots dans la phrase, etc.

7.

On ne peut distinguer en albanais que peu de traces de l'apophonie qualitative ancienne, comme par exemple *bie* < **bher-* ~ *barrë* < **bhornā* « porter » ~ « fardeau, charge », *vjer* ~ *var* < **wer-/*wor-* « suspendre, accrocher », etc. En revanche, on peut reconnaître plusieurs traces de l'apophonie quantitative dans le système verbal (cf. lat. *vĕniō/vēnī*, etc.). Dans

l'ancien parfait synthétique de l'albanais, l'ancienne apophonie quantitative s'est transformée de bonne heure en une apophonie qualitative par suite de l'évolution de **ē* et de **ā* en *o*, ainsi dans *mjel* < **mel-* ~ *mol-a* « traire », *marr* < **mers-* ~ *mor-a* « prendre », etc. Cette sorte d'apophonie, qui a cessé d'être active dès les temps préhistoriques, se rencontre à présent dans plus de 35 verbes hérités ainsi que dans leurs dérivés[4].

8.

La plupart des désinences de la déclinaison indéfinie et de la conjugaison synthétique (présent, imparfait et passé défini) sont héritées de l'indo-européen. Elles ont subi, naturellement, des changements phonétiques et analogiques. Il suffit de mentionner ici la désinence de l'ablatif pluriel *-sh* (ex. : *male-sh* « des montagnes »), qui remonte à la désinence indo-européenne **-su* du locatif pluriel. Cf. aussi les désinences du pluriel du passé défini *-më, -të, -në*, qui remontent aux désinences indo-européennes **-mes/mos, *-the, *-nti* (ex. : *pa-më, -të, -në* « nous vîmes, vous vîtes », etc.). Comme désinences nouvelles, on peut mentionner, à titre d'exemple, celles de l'imparfait en-*ja, -je, -(n)te* (*shih-ja, -je, -te* « je voyais, tu voyais », etc.).

9.

Les documents albanais ont conservé plusieurs traces du système des trois genres attesté dans les autres langues indo-européennes. Mais il faut souligner qu'en albanais, comme dans la plupart des langues indo-européennes, le neutre hérité s'est intégré graduellement aux autres genres particulièrement au genre masculin. En outre, il faut relever que le neutre attesté de l'albanais n'a de forme propre qu'au nom.-acc. singulier de la flexion définie : ainsi *djathë-të* « le fromage », *miell-të* « la farine » (mais à présent plutôt masc. *djath-i, miell-i*), etc.

---

4. V. *op. cit.* (n. 1), p. 90 *sq.*

10.

L'albanais a conservé aussi un nombre considérable de mots hérités de l'indo-européen, qui ont subi naturellement des changements phonétiques et sémantiques au cours de l'évolution préhistorique et historique de la langue. Tels sont, outre les verbes apophoniques (voir § 7) et les verbes supplétifs du type de *shoh* ~ *pa-sh-ë* « voir », etc., de nombreux autres mots : verbes, substantifs, adjectifs, prépositions, etc. Exemples : *di* « savoir », *(v)des* « mourir », *fle* « dormir », *hyj* « entrer », *korr* « récolter », *njoh* « connaître », *puth* « embrasser », *ruaj* « garder », etc. ; – *bir, bijë* « fils, fille », *dhëndër* « beau-fils », *nuse* « belle-fille », *burrë* « homme, mari », *njeri* « homme », *grua* « femme », *vjehërr* « beau-père », *ballë* « front », *gjak* « sang », *gji* « sein », *mjekër* « barbe, menton », *sy* « œil », *vesh* « oreille », etc. ; – *i athët* « acide », *i bardhë* « blanc », *i gjallë* « vif », *i gjerë* « large », *i zi* « noir », etc., – *një* « un, une », *dy* « deux », *tre/tri* « trois », etc., – *në* « à », *mbi* « sur », *me* « avec », *për* « pour », *prej* « de », etc. [5].

11.

L'albanais a suivi une évolution partiellement semblable à celle des autres langues indo-européennes, en particulier les langues balkaniques, mais sur certains points il n'y a pas de parallélisme. Les phénomènes les plus significatifs de cette évolution peuvent être présentés brièvement comme suit.

L'albanais se caractérise par un accent plutôt dynamique, qui s'est fixé de bonne heure sur une certaine syllabe du radical. Dans le système verbal il s'est fixé régulièrement sur la dernière syllabe, tandis que dans le système nominal il s'est fixé généralement sur la syllabe pénultième du radical. Mais il y a aussi plusieurs substantifs qui portent l'accent sur la première ou la dernière syllabe d'un radical polysyllabique. On comparera *flútur, flútur-a, flútur-a-ve*, etc. « papillon, papillons, aux papillons » ; *punój, punó-jmë, punó-va*, etc. « je travaille, nous travaillons, je travaillai », etc.

La fixation de l'accent sur une syllabe du radical est un phénomène achevé plusieurs siècles avant l'émigration des Albanais en Grèce (XIVᵉ siècle) et en Italie (XV-XVIᵉ siècle). Selon toute vraisemblance, c'est un

5. V. Eqrem Çabej, *Studime etimologjike në fushë të shqipes*, I (Tirana, 1982).

des phénomènes linguistiques qui ont entraîné la transformation graduelle de l'illyrien en albanais.

12.

Si l'on compare le système phonétique de l'albanais au système primitif indo-européen, on peut observer plusieurs innovations qui datent, pour une part, d'une période très ancienne. Il s'agit de changements comme l'élimination de l'ancienne opposition entre voyelles brèves et voyelles longues, l'évolution des tectales et des consonnes aspirées. etc.

L'ancienne opposition voyelles brèves ~ voyelles longues a été éliminée de bonne heure du fait surtout de l'évolution des voyelles longues. On tient pour acquis que l'ancien **ā* a donné en albanais *o* (ex. : *motër* « sœur » < **mātēr*). On a démontré aussi que l'**ē* avait évolué à date ancienne en *o*, l'**ō* en *e*, l'**ū* en *y*, etc. Ex. : *dorë* « main »< **ĝhērā*, *ne* « nous » (acc.) < **nōs*, *dyllë* « cire » < **ghūd-los*, etc.

Il faut souligner aussi que même certaines voyelles brèves ont subi de profonds changements, ainsi **o* > *a* (*natë* « nuit », cf. lat. *nox, noctis*), **e* > *ie, ia*, etc.[6].

Quelques-uns des changements que l'on vient de mentionner ont affecté même la vieille couche des emprunts au grec ancien et au latin, ainsi *mokënë/mokërë* (voir § 4), *herë* « fois » (< lat. *hōra*), *gjymtyrë* « membre » (< lat. *junctūra*), etc.

13.

En ce qui concerne l'évolution des tectales indo-européennes en albanais, on a généralement admis la thèse de H. Pedersen[7] selon laquelle les tectales palatales **k̂*, **ĝ* ont abouti à *th*, *dh* / *d* par un intermédiaire **ć, *ǵ*, et les labiovélaires **k^{w}*, **g^{w}* à *s, z* par un intermédiaire **ć, dź* > *s, z* devant les voyelles *i* et *e*, tandis que dans d'autres positions elles ont abouti à des vélaires simples. Exemples :

6. V. Gustav Meyer, *Albanesische Studien*, III (Vienne, 1892), p. 79 *sq*.
7. V. Holger Pedersen, *KZ* XXXVI (1900), 277-340.

a) vélaires indo-européennes : _ka_ «bœuf»< _*kav-_, _gjak_ «sang»< _*sokos_, _gardh_ «haie»< _*gardos_ < _*ghordhos_, _shteg_ «sentier»< _*stoighos_, etc. ;

b) palatales : _bathë_ «fève»< _*bhak̂ā_, _i athët_ < _*ak̂idus_, _derr_ < _*ĝhoiro-_, _dorë_ < _*ĝhērā_, _dhandër/dhëndër_ < _*ĝent-_, _i bardhë_ «blanc» < _*borĝhos_, etc. ;

c) labiovélaires : _pesë_ «cinq» < _*penk_^{w}_e_, _si_ «comme» < _*k_^{w}_ei_, _darsmë_ «noce» < _*dork_^{w}_i-mā_, mais _darkë_ «repas du soir», _ulk_ «loup» (cf. gr. λύκος), etc.[8].

14.

La désaspiration des anciennes occlusives aspirées s'est elle aussi achevée de bonne heure, avant l'introduction des emprunts au grec ancien et au latin : comparer _bie(r)_ < _*bherō_ avec le grec φέρω et le latin _ferō_.

15.

En ce qui concerne les phénomènes morphonologiques qui constituent la flexion interne de l'albanais, on doit relever, outre l'apophonie (v. § 7), la métaphonie et le changement de _-e/-ie-_ en _-i-_ (généralement) devant deux consonnes suivies de _-i_, ainsi : _dash_ ~ _desh_ «bélier ~ béliers», _dal_ ~ _del_ ~ _dil-ni_ «je sors, tu sors, vous sortez», etc. Ces deux phénomènes sont essentiellement le résultat d'une assimilation régressive. La métaphonie a été provoquée par un _-i_ final qui dans certains cas a laissé aussi d'autres traces, comme par exemple la palatalisation de _-k_ en _-q(= /c/)_ dans le pluriel de certains substantifs anciens comme _plak_ ~ _pleq_ < _*plak-i_ «vieillard», etc. Ces phénomènes ont eu lieu avant la pénétration des emprunts slaves en albanais, mais on en compte aussi des exemples dus à l'analogie et de date relativement récente, comme _turk_ ~ _turq_ «turc ~ turcs», _çek_ ~ _çeq-e_ «chèque ~ chèques», etc.

8. V. Shaban Demiraj, _Gjuha shqipe..._ (_op. cit._ n. 2), p. 38 _sq._

16.

Les innovations qui se sont produites dans le système morphologique de l'albanais ont eu lieu, elles aussi, dès la période prélittéraire de cette langue. Telle est par exemple la création du thème de pluriel des substantifs, qui se forme ou par l'adjonction d'une désinence (monosémantique), ou par des phénomènes phonétiques, ou bien par deux procédés réunis. Les désinences casuelles s'ajoutent à ce thème de pluriel : ainsi *mal* ~ *mal-e* ~ *male-ve, male-sh* « mont ~ monts ~ aux monts, des monts », *plak* ~ *pleq* ~ *pleq-ve, pleq-sh* « vieillard ~ vieillards ~ aux vieillards, des vieillards », etc.

Pareille innovation s'est produite aussi en roumain et en allemand, mais il s'agit de développements parallèles.

17.

Une autre innovation, et de plus grande portée, est la création d'un système complexe d'articles, qui a provoqué une série de transformations dans le système nominal de l'albanais, à savoir l'apparition non seulement de l'article postposé (qui a donné lieu à la création de la déclinaison définie), mais aussi de ce qu'on appelle l'article de liaison antéposé. Ce dernier phénomène a donné lieu, entre autres, à la création d'une classe d'adjectifs qui s'oppose à une autre classe, dénuée de cet article (ex. : *i mirë* ~ *trim* « bon ~ brave »). Cet article a servi aussi à distinguer le génitif du datif (ex. : *malit* ~ *i malit* « mont ~ du mont »). Dans les deux cas l'article de liaison antéposé se décline régulièrement, donnant lieu ainsi à l'apparition d'une nouvelle sorte de flexion au début du mot, ce qui est une grande innovation par rapport au type flexionnel indo-européen.

L'article postposé, qui s'est développé aussi en roumain, en bulgare et en macédonien, s'est agglutiné en albanais, selon toute vraisemblance, dès les premiers siècles de notre ère. L'article préposé s'est développé plus tard, mais en tout cas avant la pénétration des emprunts au slave. On ne peut donc pas établir une liaison directe entre l'article préposé de l'albanais et les deux sortes d'articles préposés du roumain, qui sont apparues plus tard [9].

---

9. V. Shaban Demiraj, *Grammatikë historike...* (*op. cit.* n. 1), p. 326 *sq.*

18.

Dans le système verbal aussi on peut constater plusieurs innovations intéressantes. Il suffit de mentionner ici les formes analytiques du parfait et du futur, les formes particulières de l'admiratif et les formes impersonnelles analytiques.

Les formes analytiques du parfait, qui se sont développées non seulement au mode indicatif mais aussi au subjonctif, à l'optatif et à l'admiratif, sont constituées par les verbes auxiliaires «avoir» et «être», comme dans les langues romanes, et le participe passé invariable en cet emploi, comme dans les langues germaniques. Exemple : *kam larë* ~ *jam larë* «j'ai lavé ~ je me suis lavé/j'ai été lavé», *kisha larë* ~ *isha larë* «j'avais lavé ~ je m'étais lavé/j'avais été lavé». Ces formes analytiques datent de la période prélittéraire de l'albanais. Selon toute vraisemblance, il s'agit d'un développement interne de cette langue [10].

19.

Quant au futur, l'albanais en connaît deux types. L'un est de type balkanique, formé avec l'auxiliaire «vouloir» (figé en albanais comme en grec) et le subjonctif (ex. : *do të bëj* «je ferai», cf. grec θα κάνω). L'autre type est formé au moyen des formes personnelles de l'auxiliaire *kam* «avoir» et des formes impersonnelles du type de *me ba* «faire» (en guègue) ou bien *për të bërë* «faire» (dans les deux dialectes). Ex. : *kam* (*ke, ka*, etc.) *me ba/për të bërë* «je ferai, tu ferais, il fera», etc. Les deux types remontent l'un et l'autre à la période prélittéraire de l'albanais [11].

---

10. *Ibid.*, p. 797 *sq.*
11. *Ibid.*, p. 822 *sq.*

20.

Il s'est développé en albanais une forme verbale particulière, qui consiste en une agglutination du parfait renversé et qui a pris le sens de l'admiratif présent. Il s'agit de la forme du type de *qen-kam* (< **qenë kam*) « on dit que je suis ». Sur ce modèle sont formés un imparfait (*qen-kësha* « on disait que j'étais »), un parfait (*paskam qenë* « on dit que j'ai été ») et un plus-que-parfait (*paskësha qenë* « on disait que j'avais été »). L'admiratif de l'albanais remonte à la période prélittéraire de cette langue.

21.

L'albanais attesté ne connaît pas d'infinitif synthétique, et l'on a beaucoup discuté la question de savoir s'il en a existé un antérieurement. Actuellement on emploie avec la valeur d'infinitif des formes analytiques, du type de *me ba* en guègue et de *për të bërë* dans les deux dialectes. La première date de la période prélittéraire, tandis que la seconde est une innovation postérieure aux premiers documents.

22.

Dans le domaine de la syntaxe, il faut relever surtout la reprise des compléments d'objet par les formes atones des pronoms personnels, comme dans la plupart des langues balkaniques. En albanais, ce phénomène date de la période prélittéraire [12].

Quant à l'ordre des mots, il faut souligner que l'albanais n'a conservé que partiellement l'ancienne liberté des langues indo-européennes. Il faut relever aussi qu'en albanais les adjectifs, les possessifs et d'autres déterminants du même type se placent régulièrement après le substantif.

12.. *Ibid.*, p. 581 *sq.*

23.

De ce qui vient d'être présenté succinctement, on peut conclure que l'albanais appartient au groupe *satəm* des langues indo-européennes, mais qu'il occupe une place à part dans ce groupe (voir § 13). Selon toute vraisemblance, il s'est développé par une évolution de l'ancien illyrien du sud entre les cinquième et sixième siècles. Au cours de son évolution, il s'est enrichi de mots nouveaux ainsi que de plusieurs formes analytiques qui se sont ajoutées aux formes synthétiques héritées. Ainsi l'albanais s'est-il transformé en une langue synthético-analytique de type balkanique.

CHAPITRE 12

LE BALTE ET LE SLAVE

Heinz Dieter POHL

1. INTRODUCTION

Les langues baltes et slaves appartiennent avec les langues germaniques à l'aire Nord-Ouest du proto-indo-européen. D'où une série de caractères communs à ces trois langues, comme les finales casuelles en *-m-* (v.sl. *vlŭkomŭ*, lit. *vilkàms*, got. *wulfam* « aux loups ») en regard des finales en *-bh-* des autres langues (skr. *vr̥kebhyaḥ*, lat. *hominibus*) ; ou la flexion de thèmes en **-yo-* des participes en **-nt-*, mais qui est moins probante, car elle peut résulter de développements parallèles indépendants : il subsiste des formes sans **-yo-*, et le phénomène n'apparaît pas dans la plus ancienne hydronymie. Les caractères communs ne sont pas répartis également entre les trois langues ; en tout cas le balte et le slave sont plus proches l'un de l'autre que de l'une des autres langues indo-européennes apparentées. Considérons d'abord les langues baltes et slaves.

1.1. Les langues baltes

On entend par « balte » un groupe de langues indo-européennes, qui ne sont plus représentées aujourd'hui que par deux langues, le lituanien et le letton. Le lituanien est à maint égard une des langues indo-européennes les plus archaïques et par là l'une des « langues clés » pour la reconstruction du proto-indo-européen. Il est parlé par environ 2 600 000 personnes en Lituanie, ancienne république socialiste de l'URSS, où les Lituaniens représentent à peu près 80% de la population. Le letton est la petite sœur du lituanien, avec lequel il représente le rameau oriental des langues baltiques; il est parlé par 1 400 000 locuteurs, soit 52% de la population de Lettonie.

Le lituanien se divise en deux groupes de dialectes : l'un occidental (bas-lituanien = žemaitique), l'autre oriental (haut-lituanien = aukštaitique); la langue littéraire se fonde essentiellement sur l'aukšaitique occidental). Le letton comprend trois groupes de parlers : l'un parlé (à l'Ouest) en Courlande et Livonie, le letton central qui est à la base de la langue littéraire, et le haut letton.

Le principal représentant du rameau occidental (disparu) des langues baltiques est ce qu'il est convenu d'appeler vieux prussien, langue originelle de la Prusse progressivement évincée par l'allemand qui s'est éteinte au XVIIe siècle (peut-être au XVIIIe par endroits). Par suite de la distribution des anciens Prussiens en de nombreux rameaux, le vieux prussien comprenait plusieurs dialectes, par exemple les « Sudauisch » et « Jatwingisch ». A l'Est du vieux prussien se trouvait la langue de l'ancienne Courlande entre la Baltique, la Venta, l'Abava et le cours inférieur de la Nemunas (Memel). Il fut refoulé au nord par le letton, au sud par le lituanien. Il occupe une position intermédiaire entre le baltique occidental et le baltique oriental, dans la mesure où les gloses, les toponymes et les emprunts aux autres langues (letton, live) permettent de le voir. Meilleure est la transmission du vieux prussien, grâce au « vocabulaire d'Elbing » (glossaire allemand-vieux prussien d'environ 800 substantifs et adjectifs), ainsi qu'aux trois catéchismes imprimés à Königsberg (I et II en 1545, III, l'« Enchiridion » en 1561, d'Abel Will). S'y ajoutent quelques moindres documents. Aussi le vieux prussien peut-il être considéré comme langue à corpus restreint; mais les noms propres, anthroponymes, toponymes, noms de fleuves, etc., ont une grande importance.

Les principales particularités des langues baltiques sont :

1) un système verbal particulier;

2) des prétérits en *-ē-* et *-ā-* (*e.g.* lit. *dìrbo* « je travaillais »; *mókė* « j'enseignais », de *dìrbti*, *mókyti*; pr. *kūra* « je construisais », *ismigē* « il mourut »);

3) une forme identique pour la 3e personne de tous les nombres (étymologiquement 3e sg., **-t* dans les thèmes plurisyllabiques, qui s'amuït

ensuite, et **-ti*, désinence primaire, dans les athématiques, cf. v.lit. *ẽsti*, pr. *ast* « est », lit. *tùri*, pr. *turri* « a ») ;

4) la 1re sg. athématique a la désinence **-mái* (*e.g.* pr. *asmai* « je suis » ; lit. *-mi*, réfléchi *-mie-s*) ;

5) la généralisation de la voyelle thématique *-a-* (< indo-européen *-o-*) ;

6) des substantifs thèmes en *-ē-* (*e.g.* lit. *žẽmė*, pr. *semmē* « terre ») ;

7) la tendance à l'épenthèse d'un *-k-* devant un groupe consonantique commençant par *s* (*e.g.* lit. *krìkštyti* « baptiser », pr. *Cricstus [-kst-]*, de *Christus*, lit. *pir̃štas* « doigt » en regard de lett. *pìrksts*) ;

8) de nombreux procédés de formation communs (nom, verbe) ;

9) des anthroponymes communs (*e.g.* lit. *Al̃girdas*, pr. *Algard*) ;

10) un vocabulaire baltique spécial (*e.g.* lit. *dangùs*, pr. *dangus* « ciel », lit. *gim̃ti*, lett. *dzìmt* « naître », pr. *gemmons* « né »).

Les langues baltiques sont séparées par d'importantes différences dialectales, ce que montrent clairement des distinctions entre le vieux prussien (baltique occidental) et le letto-lituanien (baltique oriental) :

1) gén. sg. lit. *diẽvo* (de l'ablatif indo-européen)/pr. *deiwas* (génitif indo-européen), cf. *3.3.* ;

2) lit. *-m-*/pr. *-sm-* dans la flexion pronominale, cf. *3.3.* ;

3) participe présent passif lit. *-mas* (*e.g.* lit. *nẽšamas* « étant porté » comme v.sl. *nesomŭ*, à parallèles anatoliens, louv. *-m(m)a-*, hitt.-hiér. *-ma-*)/pr. *-manas* (*e.g. poklausīmanas* « étant entendu », correspondant à gr. *-menos*, skr. *-māna-*) ;

4) le nom de nombre « neuf », cf. *3.3.* ;

5) le nom de nombre « troisième », cf. *3.3.* ;

6) les adjectifs possessifs lit. *mãnas* « mon », *tãvas* « ton », *sãvas* « son »/pr. *mais, twais, swais*, cf. *3.2.* ;

7) prétérit lit.-lett. *bi-*/pr. *bēi, be* « était », cf. *3.2.* ;

8) présents en nasale de **sed-* « être assis » et **stā-* « être debout » en vieux prussien (pr. *sindats, syndens, postanai*) en regard de lit. *sė́da, stója*, cf. *3.2.* ;

9) en vieux prussien indo-européen **ei* subsiste en toute position (*e.g.* pr. *deiwas*, ci-dessus sous 1) ;

10) le vieux prussien a des restes de neutres (en baltique oriental les anciens neutres sont devenus masculins, *e.g.* lit. *ẽžeras*/pr. *asseran*, v.sl. *jezero* « lac ») ;

11) datif pluriel pr. *-mans*/v. lit. *-mus* (mod. *-ms*) ;

12) différences dans le vocabulaire, ainsi que dans les pronoms et le comparatif.

Les principales différences entre le letton et le lituanien sont (dans la mesure où elles concernent également leurs rapports au slave) :

1) en letton *k* et *g* devant des voyelles antérieures et *j* deviennent *c* et *dz* (palatalisation, *e.g.* lett. *cits* « autre », *dzîvs* « vivant »/lit. *kìtas, gývas* ; le slave connaît un phénomène semblable, cf. *1.2.*) ;

2) les correspondants des palatales indo-européens *k' g'* sont en lituanien *š, ž,* en lette (ainsi qu'en vieux prussien et slave) *s z* (*e.g.* indo-européen **deḱm̥* « dix », lit. *dẽšimt*, lett. *desmit*, v.sl. *desętĭ*; indo-européen **ĝhei-m°*- « hiver », lit. *žiemà*, lett. *ziema*, v.sl. *zima*). Le proto-balte avait *š ž*, ce que montrent les emprunts baltiques en finnois, *e.g.* à lit. *s* correspond finn. *s.*, mais à lit. *š* répond *h*, ce qui montre que les dialectes baltiques avec lesquels le finnois est entré en contact distinguaient entre */s/* et */š/* (*e.g.* finn. *salo* « île », lit. *salà* ; finn. *heinä* « foin », lit. *šiẽnas*).

3) les noms de nombre de 11 à 19 sont en letton comme en slave du type « un, deux plus dix » (lett. *vien-pa-dsmit*, v.sl. *jedinŭ na desęte*), tandis que le lituanien forme ces numéraux à l'aide d'un élément *-lika* (*vienúolika, dvýlika*, etc.) jusqu'à 19 ; cf. germ. de même « 11 » et « 12 », got. *ainlif, twalif.*

Il existe certes d'autres différences ; en général le letton représente un état de langue plus récent que le lituanien.

1.2. Les langues slaves

Le slave fut à l'origine parlé sur une aire beaucoup plus étroite qu'aujourd'hui. Le « berceau » des Slaves se trouvait au Nord de l'arc des Carpates dans les actuelles Pologne et Ukraine occidentale ; le cœur de leur domaine était les pentes septentrionales des Carpates approximativement entre Zakopane à l'Ouest et la Bukowina à l'Est. L'expansion des Slaves dans leurs habitats historiques a dû être très rapide : le VIe et VIIe siècle de notre ère. Dans de nombreux domaines se mêlèrent des populations slaves et non slaves, et l'élément slave ne put se maintenir partout (par exemple en Grèce) ou fut trop faible pour s'imposer (par exemple en Autriche orientale) ou encore fut recouvert par d'autres populations (par exemple les Hongrois et les Allemands orientaux).

La langue slave originelle (« proto-slave ») était assez unitaire. Au contraire de ce qui se passe pour le balte, pour rendre compte de la place du slave dans les langues indo-européennes, il faut partir de l'idée que le slave dans son ensemble manifeste les mêmes affinités linguistiques. L'émiettement d'un protoslave relativement unitaire découle de l'expansion des Slaves et produit la nouvelle phase du « slave commun » (VIIIe-XIe siècle). D'un point de vue géographique apparaissent au XIIe siècle trois groupes de langues : le slave occidental (polonais, kachoube, tchèque, slovaque, sorabe et quelques langues éteintes comme le polabe), le slave

oriental ([grand)] russe, biélorusse et ukrainien), et le slave du Sud (slovène, croate ou serbe [« serbocroate » dans la terminologie scientifique], bulgaro-macédonien). Cette tripartition ne résiste pas à l'examen linguistique ; ni le slave occidental ni le slave du Sud ne représentent un type homogène. Aussi y a-t-il eu ces dernières années des tentatives pour lutter contre cette tripartition. Mais pour l'appréciation des rapports entre le slave et le balte la division du slave est sans importance.

Les principales caractéristiques du proto-slave sont :

1) disparition des voyelles arrondies (il ne reste qu'une diphtongue en *-u-*, *-au-*, qui passe ensuite à *u*). Le système vocalique proto-slave était (selon F.V. Mareš) :

hautes	ÿ̄	ȳ		i	y
	ÿ̆	y̆	=	ĭ	ŭ
basses	ä̆	ă		e	ă
	ǟ	ā		ě	a
(palatales)		(vélaires)			

La notation très répandue *o* pour *ă* vient d'une valeur phonétique ultérieure ; elle n'est pas justifiée pour le proto-slave. D'une part les plus anciens noms propres slaves qui ont *o* sont encore transmis avec *a*, *e.g.* *Dabramuzli* = *Dobromyslĭ*; de même l'ethnique *slověne* « Slaves » apparaît dans les sources latines et grecques sous la forme *Sclaveni, Sklavenoi*, etc. ; de plus l'*ă* étranger des mots d'emprunts est représenté par *o* en slave, *e.g.* lat. *păgānus* = v.sl. *poganŭ*, mais l'*o* étranger par *ŭ* (bref) ou bien par *y* ou *u* (long), *e.g.* rom. *(h)ortus* sl. *vŭrtŭ*, rom. *crōce* sl. *kryžĭ*, rom. *ratiōne* scr. *račun* « calcul ». – À côté de *au* existait aussi *jau*, noté aussi *'au* par de nombreux savants ; celui-ci donna ensuite *ju* (écrit aussi *ü*) ; il vient de l'indo-européen **eu*. La diphtongue *ei* (de l'indo-européen **ei*) doit avoir très tôt évolué en *i* ; *ai* (de l'indo-européen **ai*, **oi*) s'est monophtongué relativement tard (en « *ě*^{2} », « *i*^{2} »). Comptaient également pour diphtongues les combinaisons de *ă* et *e* avec *r l m n* (*ăr*, *el*, etc.) ;

2) harmonie syllabique : en relèvent la palatalisation des vélaires devant les voyelles d'avant (on parle de palatalisation 1. et – devant *ě*^{2} *i*^{2} – de palatalisation 2.) ; la mouillure des autres consonnes également devant les voyelles d'avant ; ce qu'on appelle la yodisation (développement de consonnes + *j*) ; et quelques autres modifications. Voici quelques exemples : « palatalisation 1. » : V. sg. v.sl. *rače, bože, duše* (de proto-sl. **rak-e, *băg-e, *daux-e*) à côté du N. sg. *rakŭ* « écrevisse » ; *bogŭ* « dieu », *duxŭ* « esprit » ; « palatalisation 2. » : impér. 2e sg. v.sl. *rĭci* « dis » (de **rek-i*, proto-sl. **rek-ai*) à côté de 1re sg. v.sl. *rekǫ*, dat.-loc. sg. v.sl. *rǫcě* (de

**rǫk-ě*, proto-sl. **rănk-ai*) à côté de *rǫka* « main » ; « yodisation » : v.sl. *pišǫ* « j'écris » (de **pis-jam*) de *pĭsati*, *mǫžĭ* « homme » (de **măngjŭ-*), v.sl. *viždǫ* « je vois » (de **vidjam*), russ. *vižu*, scr. dial. (ancien) *viđu*, v.sl. *svešta* « lumière » (de **svetja*), russ. *svęča*, scr. *sveća* ; « mouillure » : un russ. *p'at'* « cinq » suppose une prononciation préhistorique *p'ent'ĭ* (v.sl. *pętĭ*) avec une mouillure de *p* et de *t* devant les voyelles claires *e* et *i*. C'est en polonais, biélorusse et russe que la mouillure est le mieux conservée, tandis que plusieurs langues n'en ont aucune trace (par exemple slovène et serbo-croate) ;

3) loi des syllabes ouvertes : en relèvent la monophtongaison des diphtongues, la naissance des nasales-voyelles et la métathèse des liquides ; auparavant il y avait déjà eu amuïssement de toutes les consonnes à la finale de mot et à l'initiale des groupes consonantiques. Voici des exemples : monophtongaison des diphtongues (cf. *duše, rĭci, rǫcě* ci-dessus) ; naissance des nasales-voyelles : acc. sg. v.sl. *rǫkǫ* « main » (de proto-sl. **rănkam*), *desętĭ* « dix » (de proto-sl. **desentĭ* > **desimti*) ; métathèse des liquides : aboutissements divers selon les langues comme v.sl. *gradŭ* « ville, citadelle » (de *gărdŭ*), *brěgŭ* « rive » (de **bergŭ*), en slave oriental « vocalisme plein » (« Vollaut », "pléophonie", russ. « polnoglasie »), *e.g.* russ. *gorod, bereg* (russ. *-grad* dans les noms de ville est emprunté à l'ancien slavon d'Église) ; chute des consonnes finales : nom. sg. nt. *telę* « veau », G. sg. *telęte* (thème *telęt-*), aor. 2.3. sg. *može* « pouvais », « pouvait » (d'un plus ancien **măges, *măget*, avec désinences secondaires *-s* pour la 2e p., *-t* pour la 3e p.).

À la fin du proto-slave et au début d'au moins une partie du slave commun (d'où procède l'ancien slavon d'Église, la plus ancienne langue slave littéraire, à fondements slaves méridionaux), le système phonologique devait être le suivant :

voyelles orales	i	y	u	voyelles nasales
	ĭ	ŭ		
	e	o		ę ǫ
	ě	a		

consonnes :	occlusives	p	t	k	t′					
		b	d	g	d′					
	fricatives	s	š	z	ž	x				
	chuintantes	c	č	(dz)						
	sonorantes	v	j	r	l	m	n	ŕ	l′	ń.

Discutée est la valeur phonétique primaire de *t′* et *d′*, réflexes de proto-sl. **tj*, **dj* : ils apparaissent dans le plus ancien slave comme *t′* ou

ḱ ou bien *d'* ou *ǵ* (ainsi dans le slave alpin précurseur du slovène : *e.g.* le nom de personne *Boruth*, *-uch* = v.sl. *borǫšt-* de proto-sl. **bărăntj-* « combattant », dans des documents du Freising *imoki* = v.sl. *imǫšti* de proto-sl. **jĭmăntj-* « ayant », dans des toponymes de Carinthie *e.g. Radweg* de **Radoviḱe* (proto-sl. **-itje*), slovène moderne *Radoviče* « village, gens de Rado » ; en macédonien la valeur *ḱ ǵ* s'est conservée jusqu'à nos jours), mais dans les documents de l'ancien slavon d'Église (vieux bulgare) écrits en cyrillique comme *št* et *žd* (cf. ci-dessus « yodisation »).

2. Les langues baltes dans le cadre des langues indo-européennes

Aussi bien le baltique que le slave entretiennent des relations de parenté diverses avec d'autres langues indo-européennes, l'un ou l'autre parfois, plus souvent tous deux ensemble (cf. *1.* en ce qui concerne le germanique).

2.1. Le balte et les autres langues indo-européennes

Les rapports de parenté du balte vont dans toutes les directions, ce que souligne bien l'hydronymie paléo-européenne : beaucoup de noms dispersés sur ce vaste domaine ont souvent un correspondant en baltique (*e.g. At(h)esis*, aujourd'hui *Etsch/Adige* en haute-Italie, cf. *Atesỹs* en Lituanie ; *Apsona*, aujourd'hui *Aisne* en France, cf. *Apsuonà* en Lituanie ; *Varia*, aujourd'hui *Vaire* en France, cf. *Vãrė* en Lituanie ; *Dravus*, aujourd'hui *Drau/Drava*, en Autriche et dans l'ex-Yougoslavie, cf. *Drawe* en Prusse orientale, etc. ; toutes ces données d'après W.P. Schmid).

Si l'on fait abstraction des rapports étroits entre balte et slave, le balte partage une série d'isoglosses avec le germanique et l'indo-iranien, mais aussi avec les Balkans et le Nord de l'Adriatique, à quoi s'ajoutent des relations récentes d'emprunt avec les langues finnoises. Deux isoglosses balto-germaniques importantes sont l'ordinal « premier », v.a. *forma-*, got. *fruma-*, lit. *pìrmas*, pr. *pirmas* (d'indo-européen **pr̥H-mo-* en regard de **pr̥H-wo-* > skr. *pū́rvaḥ*, sl. *prŭvŭ*) et les noms de nombre « 11 » et « 12 » (cf. *1.1*). Deux isoglosses caractéristiques partagées avec l'indo-iranien sont entre autres les voyelles longues de lit. *výras*, skr. *vīrá-* « homme » (en regard *e.g.* de lat. *uir*, germ. *wer-*) et de lit. *vė́jas*, skr. *vāyú-* « vent » (en regard de v.sl. *vĕtrŭ* comme de lit. *vė̃tra* « tempête » ou du plus récent lat. *uentus*).

2.2. Le slave et les autres langues indo-européennes

Les rapports les plus étroits se trouvent entre le slave et le balte, et entre ces deux derniers ensemble à nouveau avec le germanique et l'indo-iranien. Il existe en outre des contacts et des rapports étroits avec l'iranien, mais aussi l'italique, l'albanais, l'anatolien et l'arménien, pour ne nommer que les plus importants.

Avec l'iranien le slave a des correspondances aussi bien grammaticales que lexicales, *e.g.* instr. sg. des thèmes en *-ā-*, v.sl. *-oją*, iran. *-ayā*, thème pronominal v.sl. *ovŭ*, iran. *ava-* «celui-ci», v.sl. *radi*, v.p. *rādiy* «à cause de», etc., ou bien au plan du vocabulaire v.sl. *bogŭ* «dieu», (en composition) «part», *e.g.* *bogatŭ* «riche» (proprement : «pourvu d'une part»), *nebogŭ* «pauvre» (proprement : «n'ayant pas de part»), v.p. *baga-* «dieu», av. *baγa-* «dieu; part» (en regard de lit. *diẽvas*, v.isl. *tivar* «dieux», skr. *devá-* «divinité», etc.), v.sl. *slovo*, av. *sravah-* «mot», v.sl. *zadŭ* «derrière», av. *zadah-* «le derrière», v.sl. *zŭlŭ* «méchant», av. *zūrah-* «injustice, tromperie» (ces exemples illustrent en même temps les accords dans le domaine des représentations religieuses). Il y a bien d'autres emprunts à l'iranien en slave faits à des époques diverses.

Dans le domaine des correspondances entre slave et latin on trouve des exemples comme *e.g.* v.sl. *mrŭtvŭ* «mort»/lat. *mortuus*; souvent le baltique participe aux isoglosses, *e.g.* v.sl. *vetŭxŭ*, lit. *vẽtušas*, lat. *vetus*; ou au sens de «corps céleste lumineux» se rencontrent v.sl. *luna* «lune», pr. *lauxnos* [-ksn-] «étoiles, constellation»/lat. *lūna* (cf. aussi av. *raoxšna-* «lumière»). Remarquables sont les concordances dans la formation des mots entre hittite et slave dans le cas de hitt. *-asti-* et *-tal(l)a-* et v.sl. *-ostĭ* et *-tel'ĭ*, *e.g.* hitt. *dalugasti-* «longueur» (de *daluki-* «long») et avec le même sens v.sl. *dlŭgostĭ* (de *dlŭgŭ*), et hitt. *aršanatalla-* «envieux» (de *aršaniia-* «envier»), v.sl. *učitel'ĭ* «maître» (de *učiti* «enseigner»). Avec l'arménien, le slave partage le participe en *-l-* tiré du thème d'aoriste.

3. Les relations entre le balte et le slave

Le rapport entre le balte et le slave est l'un des problèmes les plus discutés au sujet de la division dialectale des langues indo-européennes. Des correspondances frappantes (par exemple pour l'accent) d'une part s'opposent à des différences profondes (par exemple dans le domaine du système verbal) d'autre part. Il en résulte des opinions très divergentes;

si l'on croyait encore au XIX[e] siècle que le balte et le slave étaient les rejetons d'un « balto-slave » commun, depuis 1908, date à laquelle l'indo-européanisant français A. Meillet a pour la première fois remis en question l'unité balto-slave, la discussion n'a jamais cessé. Aujourd'hui la majorité des chercheurs croit à une unité très ancienne entre les deux langues (à une époque tardive de l'indo-européen, sans doute dans le cadre de l'aire Nord-Est de l'Europe, au III[e] millinéaire avant J.-C.). De plus vers 2000 – 1500 avant J.-C. les ancêtres des Baltes et des Slaves eurent des contacts étroits, auxquels on doit un faisceau très ancien d'isoglosses ; puis vint le temps de la séparation et des développements séparés (à partir de 1500 environ). C'est seulement plus tard que les contacts reprirent entre les deux langues (à partir du début de notre ère), d'abord entre Baltes occidentaux et Slaves, plus tard (dans la première moitié du 1[er] millinéaire après J.-C.) entre Baltes orientaux et Slaves. A partir de 600 après J.-C. (et peut-être même avant) l'expansion des Slaves recouvre le domaine balte, ce qui aboutit à une influence durable du slave sur le balte oriental (lituanien et lette).

Cette situation se reflète dans la distribution des caractères communs aux deux langues et explique en même temps les différences flagrantes entre les deux, ce à quoi seront consacrés les paragraphes suivants (*3. 1-5*).

3.1. Coïncidences anciennes et importantes entre le balte et le slave

Le balte et le slave manifestent des correspondances très anciennes dans le traitement des laryngales, l'élimination des aspirées, l'évolution des liquides et nasales syllabiques, dans la satémisation, dans l'évolution de l'*s* indo-européen après *i u r k*, dans la genèse des intonations, et dans la formation de morphèmes de dérivation identiques.

Les « laryngales » sont dans les deux langues (comme en iranien) amuïes en syllabe intérieure (en slave probablement en toute position), mais ont laissé des traces à travers les intonations ; les aspirées (*bh dh ph th*, etc.) ont disparu sans laisser de traces, en d'autres termes ont été remplacées par les occlusives non aspirées correspondantes (*b d p t*, etc.). Voici quelques exemples relatifs à la formation des mots : abstraits en *-bā-* (lit. *-yba, e.g. ganýba* « garde, pâture » de *ganýti* « paître »/v.sl. *-ība, e.g. družība* « amitié » de *drugŭ* « ami », à rapprocher de lett. *draudzība* « sodalité, amitié ») ; noms d'agent en *-tā-yo-* (développement du type gr. ποιητής, *e.g.* lit. *artójas* = v.sl. *ratai* « laboureur ») ; dans la dérivation verbale *e.g.* lit. *-áuju* = v.sl. *-ują* (cf. lit. *keliáuju* « aller son chemin », prét. 3[e] p. *keliãvo*/v.sl. *kupują* « acheter », aor. 3[e] sg. *kupova*).

Les liquides voyelles (*r̥ l̥*) et nasales voyelles (*m̥ n̥*) ont donné en baltique et (à l'origine) aussi en slave le plus souvent *ir il im in*, mais également *ur ul un um* : à une majorité de diphtongues en *i* s'oppose une minorité de diphtongues en *u*, les deux langues coïncidant dans l'ensemble jusque dans le détail (*e.g.* lit. *širdìs*, v.sl. *srĭdĭce* de proto-sl. **sĭrdĭ-kă-*, indo-européen **ḱr̥d-* « cœur », lit. *kùrti* « construire », v.sl. *krŭčii* « forgeron », c'est-à-dire proto-sl. **kŭr-* + un suffixe *-čijĭ*, d'indo-européen **kr̥-*).

La satémisation et le passage d'indo-européen *s* à *š* (qui passe ensuite en slave à *x*) après *i u r k* concernent aussi notoirement d'autres langues (entre autres l'indo-iranien) ; mais ce qui est remarquable, c'est que :

1) il y a des exceptions à la satémisation (en balte plus qu'en slave, sans que les deux s'accordent toujours, *e.g.* lit. *žąsìs* en regard de v.sl. *gǫsĭ* « oie » d'indo-européen **ǵhans*, lit. *šlovė̃* « gloire » à côté de *klausýti* « entendre », mais v.sl. *slava* et *slyšati*, d'indo-européen **ḱlew-*), et que

2) le balte n'est que partiellement affecté par le passage de *s* à *š* (*e.g.* lit. *viršùs* = v.sl. *vrŭxŭ* « haut, dessus », mais lit. *saũsas* à côté de v.sl. *suxŭ* « sec ».

Cas sans doute unique dans les langues indo-européennes, le balte et le slave ont développé des intonations sur toutes les syllabes du mot, l'opposition aigu/circonflexe, selon la même formule : les voyelles longues et les diphtongues à premier élément long (même si elles sont nées d'une diphtongue à premier élément bref ou d'une liquide + « laryngale ») reçoivent l'aigu, les diphtongues à premier élément bref le circonflexe. Parmi les langues slaves le témoin principal pour les intonations est le serbo-croate : à un aigu lituanien répond une intonation descendante brève, à un circonflexe, une intonation descendante longue (le russe, entre autres, en a des traces dans des groupes « pléophoniques »), *e.g.*

lit.	*bóba*	scr.	*bȁba* « vieille femme »
	kárvė		*krȁva* « vache » (russ. *koróva*)
	várna		*vrȁna* « corneille » (russ. *voróna*)
	kálti		*klȁti* « abattre » (russ. *kolót'*)
	žĩrnis		*zȑno* « blé »
	saũsas		*sûh* « sec »
	gar̃das		*grâd* « citadelle », ville (russ. *górod*)
	vir̃šų		*vȓh* « en haut, au-dessus de »
	žam̃bas		*zûb* « dent »
	var̃nas		*vrân* « corbeau » (russ. *vóron*).

Les deux langues montrent également d'étonnants parallélismes dans des cas d'altération d'accent, *e.g.*

lit.	*žiemà*	scr. dial.	*zīmȁ*	russ.	*zimá*	(N.sg.)
	žiemõs		*zīmé*		*zimý*	(G.sg.)
	žiẽmą		*zȋmu*		*zímu*	(A.sg.)
	žiẽmos		*zȋme*		*zímy*	(N.A.pl.)

L'ordinal « huitième » est aussi une ancienne caractéristique commune au balte et au slave : v.lit. *ãšmas*, pr. *asmus*, v.sl. *osmŭ* (en regard de lat. *octāuus*, gr.ὄγδο(ϝ)ος.) Pour le vocabulaire, voir ci-dessous *3.4.*.

3.2. Coïncidences entre le balte occidental et le slave

Les principales correspondances entre le baltique occidental et le slave sont :

1) les adjectifs possessifs : pr. *mais, twais, swais* auxquels correspondent exactement v.sl. *moi, tvoi, svoi* « mon, ton son » (en regard de lit. *mãnas, tãvas, sãvas*) ;

2) les présents à nasale des verbes indo-européens **sed-* « être assis », **stā-* « être debout » (v.sl. *sędǫ*, c'est-à-dire proto-sl. **send-, stanǫ*, pr. part. *sindats, syndens* ainsi que *po-stanai* « devient » en regard de lit. *sė́du, stóju*) ;

3) le prétérit pr. *bēi, be* et v.sl. *běxŭ, bě*, etc., « était » (en regard de lit. *bùvo*, qui est une innovation).

3.3. Coïncidences entre le balte oriental et le slave

Les principales correspondances entre le baltique oriental et le slave sont :

1) le génitif des thèmes indo-européen en *-o-* : le slave et le baltique oriental lui ont substitué l'ablatif indo-européen (i.-e. tardif **-ād/-ōd* de **-o-H_2ed-*) : lit. *vil̃ko*, v.sl. *vlŭka* « du loup » (les deux formes ne se laissent comparer que si on pose **-ād*). Le vieux prussien partage avec le germanique la désinence répandue de génitif sg. **-os(y)o/-es(y)o, e.g.* pr. *deiwas* « du dieu », got. *wulfis* « du loup » ;

2) le nom de nombre « neuf » (v.sl. *devętĭ*, lit. *devynì*, lett. *deviņi* en regard de pr. *newīnts* « neuvième » comme got. *niun* à initiale *n-*, cf. skr. *náva-*, lat. *nouem*, du type indo-européen commun ; les formes en *d-* viennent probablement d'une contamination avec le nom de nombre « dix », v.sl. *desętĭ*, lit. *dẽšimt*) ;

3) l'ordinal « troisième », lit. *trẽčias*, lett. *trešais*, v.sl. *tretĭi* (en regard de pr. *tīrts*, auquel skr. *tr̥tīyaḥ* correspond pour le radical ; on ne peut savoir si got. *þridja* repose sur **tret-* comme les formes du baltique oriental et du slave ou sur **trit-* comme gr. τρίτος ;

4) Dat.-loc. sg. *-m-* des pronoms (au lieu du *-sm-* des autres langues indo-européennes), *e.g.* dat. sg. lit. *jám, tám, šiám*, v.s.l. *jemu, tomu, semu* en regard de pr. *stesmu*, *schismu*, cf. skr. *tasmai*, got. *þamma*. Au loc.sg. répond à un lit. *tamè* un v.sl. *tomĭ*.

3.4. Différences majeures entre le balte et le slave

Une différence frappante entre baltique et slave consiste en ce que tous deux traitent pareillement **ă* et **ŏ* (comme le germanique, en *ă*, puis en slave *o* à une date relativement tardive), tandis que **ā* et **ō* coïncident en slave (en *a*, originellement long) et en germanique (en *o*), mais non en baltique, où ils restent séparés :

indo-européen **māter-* « mère », lit. *mótė* « épouse », v.sl. *mati*
indo-européen **dō-* « donner », lit. *dúo-ti*, v.sl. *da-ti*.

En outre la satémisation s'est déroulée différemment dans les deux langues (proto-balt. *š ž*, proto-sl. *s z* d'indo-européen **ḱ*, **ǵ*).

En morphologie, il y a des différences notables, entre autres les désinences verbales personnelles, aux 1[re] et 2[e] p.sg. et 3[e] p.pl. :

1.p.sg. lit. *-u* (i.-e. **-ō*)	v.sl.	*-ǫ* c'est-à-dire probablement i.-e. **-ō* + **-m*, contamination des deux désinences de 1[re] p.sg.)
2.p.sg. lit. *-i*	v.sl.	*-ši*
(tous deux d'origine contestée)		
3.p.pl. lit. *–t(i)*	v.russ.	*-ǫtĭ*, *-ętĭ* (contenant i.-e. **-nti*)

c'est-à-dire la désinence du singulier, phénomène qui n'a pas encore reçu d'explication satisfaisante, cf. 1.1).

Certes le balte et le slave coïncident en leurs prétérits en **-ā-* et **-ē-* ; cependant, en leurs structures, les flexions de ces prétérits ne peuvent guère être comparées. Il existe des concordances anciennes comme lit. *sùko* « tournais », *álko* « avais faim », v.sl. *sŭka, al(ŭ)ka*, mais en balte le thème du prétérit dépend de celui du présent, au contraire de ce qui se passe en slave, *e.g.* v.sl. *beretŭ* « recueille, lit », aor. *bĭra, kupujetŭ* « achète », aor. *kupova, viditŭ* « voit », aor. *vidě*, infin. *bĭrati*, *kupovati, viděti*, part. prét. act. *bĭralŭ, kupovalŭ, viděŭ* (*-a-* et *-ě-* sont devenus thèmes « seconds », tandis que les

correspondants baltiques ont des valeurs purement prétéritales, lit. 3ᵉ p.sg. *dìrb-o* « travaille », *mók-ė* « enseigna », 1ʳᵉ p.sg. *dìrbau, mókiau*, 1ʳᵉ p.pl. *dìrbome, mókėme* ; en slave s'ajoute au thème d'« aoriste ou thème second » le morphème d'aoriste proprement dit *-x-*, 1ʳᵉ p.sg. *bĭra-x-ŭ*, 1ʳᵉ p.pl. *bĭra-x-omŭ*, etc.). Le baltique connaît le futur en *-s-*, le slave l'aoriste en *-s-* (le plus souvent passé à *-x-*), ce qui n'a guère à susciter de commentaires.

Le balte n'a pas les correspondants des participes slaves en *-lo-* du prétérit, *-(e)no-* du passif (à côté de *-to-* comme en germanique, mais avec une distribution différente).

Il est vrai que les correspondances sont nombreuses dans le vocabulaire, *e.g.* :

« charrue »	lit.	*árklas*	v.sl.	*ralo* (proto-sl. **ărtlă-*)
« jeune »		*jáunas*		*junŭ*
« tête »		*galvà*		*glava*
« corbeau »		*var̃nas*		*vranŭ*
« corneille »		*várna*		*vrana*
« jour »		*dienà*		*dĭnĭ* ;

mais il existe aussi de graves divergences, par exemple dans la terminologie religieuse (« dieu » : v.sl. *bogŭ*, lit. *diẽvas*, cf. 2.2.), ainsi que dans le vocabulaire de base : manquent en slave les correspondants de lit. *výras* « homme », lit. *žõdis*/lett. *vārds*/ pr. *vīrds* « mot », lit. *sẽnas* « vieux », *dantìs* « dent » (le slave a *člověkŭ, slovo, starŭ, zǫbŭ*), etc. Il y a certes déshérence dans les deux langues de termes indo-européens (comme par exemple **genH₁-* « engendrer, naître », **pH₂ter-* « père », **ped-* « pied », qui n'ont laissé que des traces isolées) ; mais les innovations qui les ont remplacés diffèrent dans les deux langues (v.sl. *roditi, otĭcĭ, noga*, lit. *gim̃ti, tė́vas, kója*).

3.5. Solidarités récentes et évolutions parallèles

Le baltique et le slave ont eu maints développements communs, sans qu'il faille y reconnaître le résultat de rapports directs. Ainsi le développement des intonations a produit des résultats pour une part étonnamment comparables (cf. *3.1.*) ; mais cela ne doit pas faire perdre de vue l'apparition de divergences (aux similitudes dans le domaine nominal s'opposent les différences dans le domaine verbal). Des tendances au nivellement dans les paradigmes flexionnels comme la confusion des déclinaisons consonantique et en *-i-*, et l'introduction de **tos, tā* au nominatif sg. des démonstratifs au lieu de **so, *sā* sur le modèle des cas

obliques sont des phénomènes fréquents dans les langues indo-européennes. Même la formation d'une flexion d'adjectifs déterminés ne saurait être surestimée, car elle n'a été grammaticalisée qu'après l'entrée en vigueur des accidents phonétiques affectant en slave la finale (où ne sont pas admises les consonnes), *e.g.* nom.sg. lit. *-as* + *jis* > *-asis*, mais v.sl. *-ŭi* ou *-yi*, c'est-à-dire *-ŭ* + **jĭ*, ce qui suppose le remplacement de l'indo-européen **-ŏs* par **-ŭs* et la chute de *-s* final.

En réalité, du fait de l'extension du slave aux dépens du balte, et des contacts intensifs pluriséculaires entre balte et slave, les langues baltes et slaves ont développé des points communs qu'il n'est pas nécessaire de considérer comme héritages d'étapes linguistiques antérieures. Il en va ainsi par exemple pour la palatalisation du letton (lett. *dzîvs* en regard de lit. *gývas*, cf. v.sl. *živŭ*), la mouillure du lituanien, les noms de nombre de 11 à 19 en lette (cf. *1.1.*), l'aspect verbal, l'instrumental prédicatif, etc.

4. CONCLUSION

Le problème des rapports entre balte et slave relève aussi bien de la linguistique historique comparative (génétique) que de la linguistique aréale ; la recherche relative aux contacts linguistiques peut contribuer à la solution du problème. L'hypothèse d'une très ancienne unité du pré-balte et du pré-slave (de la période de l'indo-européen tardif : Introduction), suivie d'une séparation et de développements indépendants puis à nouveau de rapprochements tardifs du baltique et du slave, peut être étayée par :

1) des correspondances anciennes et importantes entre les deux langues (cf. 3.1) ;

2) des différences caractéristiques entre les deux (cf. *3.4*) aussi bien que des développements parallèles indépendants (cf. *3.5*) ;

3) des isoglosses entre le baltique occidental et le slave (cf. *3.2*) et le baltique oriental et le slave (cf. *3.3*) dues à un rapprochement précoce ;

4) une interpénétration due à l'expansion des Slaves à époque historique (cf. *3.5*).

Pour être complet, signalons que d'aucuns ont pensé que le slave était issu de dialectes baltiques périphériques (parfois sous l'influence d'autres langues, comme l'iranien ou l'« ancien balkanique »).

Bibliographie

BIRNBAUM H. (1970). – « Four approaches to Balto-Slavic », *Festschrift Stang* (*Donum Balticum*, Stockholm), p. 69-76.

– (1975). – *Common Slavic. Progress and problems in its reconstruction.* Cambridge/Mass.

BIRNBAUM H., MERRILL P.T. (1985). – *Recent advances in the reconstruction of Common Slavic* (1971-1982). Columbus/Ohio.

ERHART A. (1958). – « Zum Problem der baltisch-slavischen Spracheinheit », *Sborník filosofické fakulty Brněnské University* 7, A6, p. 123-130.

FRAENKEL E. (1950). – *Die baltischen Sprachen. Heidelberg.*

GIMBUTAS M. (1963). – *The Balts.* London – New York – Washington.

– (1971). – *The Slavs.* London – New York – Washington.

KARALIŪNAS S.K. (1968). – « Einige Probleme der ältesten baltischen und slavischen Sprachbeziehungen », *Lietuvių kalbotyros klausimai* 10, 7-100 (en lituanien, avec résumé en allemand).

KLIMAS A. (1970). – « Baltic, Germanic and Slavic », *Festschrift Stang* (*Donum Balticum*, Stockholm), p. 263-269.

KRONSTEINER O. (1977). – « Ist die Einteilung der slawischen Sprachen in Ost-, West- und Südslawisch gerechtfertigt ? », *Österreichische Namenforschung*, p. 17-28.

MAREŠ F.V. (1980). – « Die Tetrachotomie und doppelte Dichotomie der slavischen Sprachen », *Wiener Slavistisches Jahrbuch* 26, p. 33-45.

– (1986). – « Vom Urslavischen zum Kirchenslavischen. » *In* : Rehder P. (Hg.), *Einführung in die slavischen Sprachen* (Darmstadt) p. 1-19.

MAŽIULIS V. (1970). – Baltų ir kitų indoeuropiečių kalbų santykiai. Vilnius (« Relations entre les langues baltiques et les autres langues indo-européennes » [avec résumé en allemand p. 328-599]).

MEILLET A. (1908). – *Les Dialectes indo-européens*, Paris.

MEILLET A., VAILLANT A. (1934). – *Le Slave commun*, Paris.

POHL H.D. (1977). – « Forschungsbericht : Historische slavische Sprachwissenschaft seit 1945 », *Kratylos* 22, p. 1-39.

– (1981). – « Sprachverwandtschaftsmodelle (zur arealen Giederung des slavischen Sprachgebietes) », *Slavistična revija* 29, p. 401-411.

– (1982). – « Die balto-slavische Einheit – areal gesehen », *Wiener Slavistisches Jahrbuch* 28, p. 77-91. (Version abrégée de : « Baltisch und Slavisch », *Klagenfurter Beiträge zur Sprachwissenschaft* 6, 1980, 58-101 und 7, 1981, p. 93-126).

PORZIG W. (1954). – *Die Gliederung des indogermanischen Sprachgebietes*, Heidelberg.

SADNIK L. (1959). – *Slavische Akzentuation* I : Vorhistorische Zeit. Wiesbaden.

SCHMID W.P. (1973). « Baltische Sprachen und Völker. Baltoslawische Spracheinheit », *In : Reallexikon der germanischen Altertumskunde* II, p. 14-22.

– (1978). – *Indogermanistische Modelle und osteuropäische Frühgeschichte*, Mainz.

SENN A. (1966). – « The relationship of Baltic and Slavic », *In* : Birnbaum H., Puhvel J. (Hg.), *Ancient Indoeuropean dialects* (Berkeley-Los Angeles), p. 139-151.

– (1970). – « Baltic and Slavic linguistic relations », *Festschrift Stang* (*Donum Balticum*, Stockholm), p. 485-494.

STANG C.S., (1942). – *Das slavische und baltische Verbum*, Oslo.

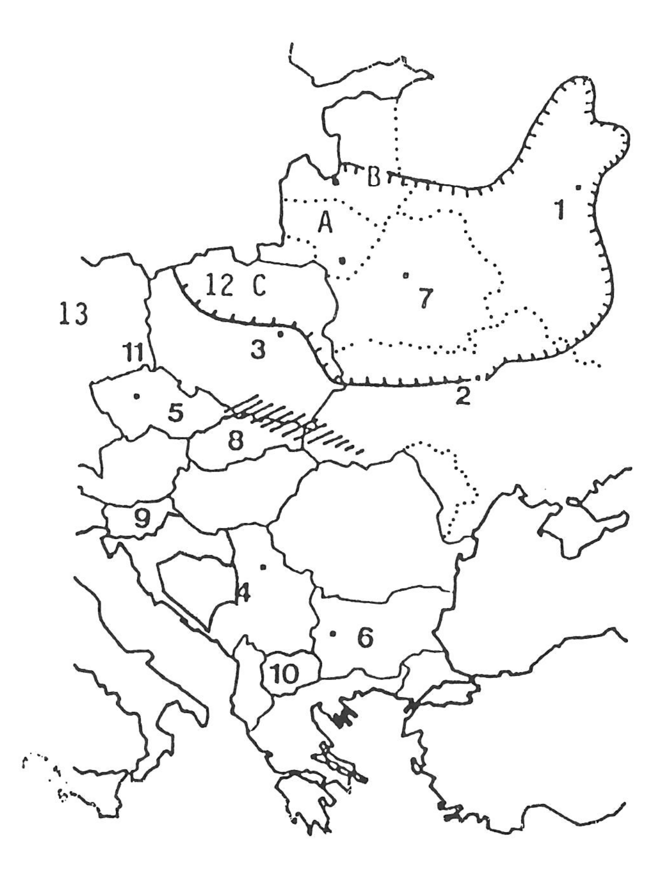
– (1957). *Slavonic accentuation*, Oslo.

– (1966). – *Vergleichende Grammatik der baltischen Sprachen*, Oslo.

TRUBAČEV O.N. (1982). – « Jazykoznanie i ètnogenez slavjan. Drevnie slavjane po dannym ètimologii i onomastiki ("La linguistique et l'éthnogenèse des Slaves. Les anciens Slaves d'après les données de l'étymologie et de l'oonomastique") ». *Voprosy jazykoznanija* 4, 10-26 und 5, 3-17.

UDOLPH J. (1979). – *Studien zu slavischen Gewässernamen und Gewässerbezeichnungen*, Heidelberg.

VAILLANT A. (1950-1977). – *Grammaire comparée des langues slaves*, Paris, 5 vol.

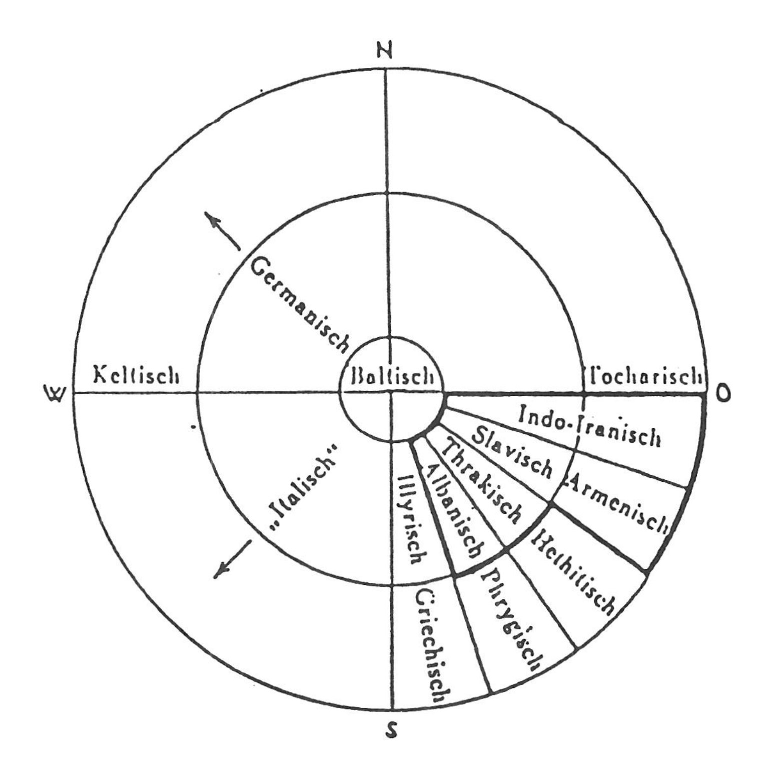


Les langues slaves et baltiques d'après leur importance quantitative
(en millions de locuteurs, † = éteintes)

1. russe (140)	7. biélo-russe (6,7)	13. polabe (†)
2. ukrainien (35)	8. slovaque (4,5)	
3. polonais (34)	9. slovène (2)	A. lituanien (2,6)
4. serbo-croate (18)	10. macédonien (1,5)	B. lette (1,4)
5. tchèque (9,7)	11. sorbe (0,05)	C. v. prussien (†)
6. bulgare (8,3)	12. kachoube (quelques milliers)	

/////// berceau présumé des Slaves (d'après J. Udolph)

ㅗㅗㅗㅗ expansion préhistorique maximale des Baltes au témoignage des noms de fleuves d'après M. Gimbutas)



La position centrale du baltique dans l'*Indogermania* (d'après W.P. Schmid) : le baltique a des rapports aussi bien avec l'Ouest qu'avec l'Est; en étoile sont dans le premier cercle des langues qui entretiennent les rapports les plus étroits avec le baltique, dans le second cercle (à l'extérieur) prennent place les autres langues. La ligne en caractères gras souligne l'isoglosse «satem».

CHAPITRE 13

GERMANIC (LE GERMANIQUE)

Jay H. JASANOFF

1. INTRODUCTION

The Germanic languages of the present day occupy a large expanse of northern and central Europe, including Scandinavia and the British Isles; one language of the family, English, has spread in recent times to cover the greater part of North America, most of Australia, and smaller areas elsewhere. But like Romance, the other Indo-European language group with a worldwide distribution, Germanic started from insignificant beginnings in an area remote from the centers of European civilization. Early in the first millennium B.C., Proto-Germanic, the source of the modern languages, was spoken by a small population of Iron Age agriculturalists in southern Scandinavia and the adjacent Baltic and North Sea coasts. In the centuries that followed, Germanic tribes slowly spread to the south and southeast, largely displacing or absorbing the Continental Celts, who had occupied much of central and eastern Europe a half millennium earlier. Before the beginning of the Christian era, Germanic peoples stood across the Rhine

from the fields of Roman Gaul, and intermittently menaced the prosperous Greek settlements on the shores of the Black Sea.

The culture of the ancient Germans, at the moment of their first appearance in history, was distinctive and homogeneous. Germanic society was characterized by a shared religion, a shared legal and political system, a shared *ars poetica* and, of immediate interest to us here, a shared language. Interesting glimpses of this world are afforded by the *Germania* of Tacitus, written around the year 100 A.D. By that time, the Germanic dialects had been diverging for three or four centuries. The tribes which, starting from what is now western Poland, had moved toward the Danube basin and the steppes of southern Russia, constituted an incipient *East Germanic* linguistic group, which differed, only slightly at first, from the *North Germanic* of southern Scandinavia and the *West Germanic* of Jutland, Germany proper and the non-Celtic part of the Low Countries. The latter group was itself an aggregate of closely related dialects, of which modern scholars often recognize three: *North Sea Germanic* or *Ingvaeonic, Rhine-Weser Germanic* and *Elbe Germanic*, corresponding to tribal divisions recorded by Tacitus and other Roman authors. East Germanic, now extinct, gave rise to *Gothic*, the language of the Arian bishop Wulfila's translation of the Bible; the substantial surviving fragments of this fourth century text constitute our earliest extensive remains of any Germanic language. A late form of Gothic is believed to persist in *Crimean Gothic*, the dialect of a short sixteenth-century word list. North Germanic is first attested in short inscriptions in the Runic alphabet (cf. PGmc. **rūnō* « secret », several of which go back to the beginning of the third century or even earlier. The language of these meager texts is usually termed *Primitive Norse (Urnordisch)*; until the sixth century it was more conservative than Gothic, and hardly distinguishable, in phonology at least, from Proto-Germanic itself. The earliest North Germanic literary languages, often collectively known as *Old Norse*, are *Old Icelandic* and *Old Norwegian*, together constituting *West Norse*, and *Old Danish* and *Old Swedish*, together constituting *East Norse*. All are attested from the twelfth century; *Modern Icelandic, Norwegian, Danish* and *Swedish* (from c. 1500) are their direct descendants [1]. The present-day languages of the Scandinavian mainland are structurally quite different from Old Norse, having participated for nearly a thousand years in the general western European drift toward morphological simplification. Icelandic, on the other hand, has changed but little from its medieval form, and is by far the most archaic Germanic language still spoken.

The West Germanic languages present a more complicated picture. The Ingvaeonic group is best represented in the early medieval period

1. Another descendant of Old Norwegian is *Faroese*, spoken by about 43,000 people in the Faroe Islands.

by *Old English*, formerly called *Anglo-Saxon*, which was brought to Britain, apparently amidst scenes of great violence, by invaders from the Continent in the fifth and sixth centuries. Old English is attested from the eighth century; after c. 1100 it is usual to speak of *Middle English*, and after c.1450 of *Modern* or *New English.* Closely related to English is *Frisian*, the earliest documents of which, in *Old Frisian*, date from the thirteenth century. *Old Saxon*, attested from the early ninth century, and its descendants *Middle* and *Modern Low German*, are mixed dialects, with an Ingvaeonic core overlaid by elements from the neighboring non-Ingvaeonic speech of the Franks. The Franks, a tribal confederation first mentioned in the third century, were the principal bearers of Rhine-Weser Germanic. Their northernmost dialect, poorly attested at the end of the tenth century as *Old Low Franconian*, became the basis of *Middle* and *Modern Dutch* (from c. 1500) [2]. The remaining Frankish dialects, *Middle Franconian, Rhenish Franconian*, and *East Franconian*, collectively constitute the "Middle German" (*mitteldeutsch*) component of *High German*, or German proper. *Old High German*, documented from c. 750, is an assemblage of Middle German dialects and dialects of the "Upper German" (*oberdeutsch*) type, the latter consisting chiefly of *Bavarian* and *Alemannic*, representing the old Elbe Germanic group. Around 1050 Old High German reached the stage known as *Middle High German*, which in turn developed into *Modern* (or *New*) *High German* around 1350. Modern literary German is based primarily on East Franconian, the dialect of Martin Luther's sixteenth-century translation of the Bible. Another descendant of Old High German is Judeo-German or *Yiddish*, with a history extending back to the eleventh century.

The map of the Germanic languages is thus, from the beginning of the historical record, crisscrossed by dialect divisions divisions whose interpretation and significance have traditionally been a source of controversy. The old view that North and East Germanic are more closely related to each other than to West Germanic is no longer widely held. The special points of resemblance between Gothic and Old Norse are either common retentions, like the ending *-t* in the 2 sg. preterite (cf. Go., ON *namt* vs. OHG *nāmi* «you took», or structurally insignificant innovations, like the occlusion of **ww-* and **-jj-* to **-ggw-* and **ddj-/*-ggj-* (cf. Go. *twaddje*, ON *tveggja* vs. OHG *zweiio* «two [gen. pl.]»). More impressive are the innovations common to North and West Germanic, such as the change of **ē* to **ā* (cf. ON *sáð*, OHG *sát* vs. Go. *seþs* «seed», the development of an *s*-extended demonstrative with the meaning «this» (cf. ON *þessi*, OHG *desēr*), and the creation of a seventh class of ablauting strong verbs (cf. ON *heita*, pret. *hét*, OHG *heizan*, pret. *hiaz* vs. Go. *haitan*, pret. *haihait* «call»). An-

2. Here too belongs *Afrikaans*, spoken in the Republic of South Africa by the descendants of the original Dutch settlers.

other debate centers about the status of West Germanic itself, which some scholars have refused to recognize as a historically valid linguistic grouping. Such doubts are probably exaggerated. While Old English, Old Saxon and Old High German are far less alike than, say, Old Icelandic and Old Swedish, they show too many common developments (e.g., the change of **ϑ* to **d*, the gemination of consonants before **-j-*, the loss of final **-z* after unstressed vowels) for the assumption of a Common West Germanic period to be dispensed with entirely. West Germanic is best thought of as a cluster of closely related dialects, with the deepest isoglosses separating Ingvaeonic from the others.

The contentious state of Germanic dialectology makes it easy to lose sight of the fact that the early Germanic languages are all very similar. The Germanic peoples of the early Middle Ages were probably in some sense aware of their kinship. Their traditional heroic literature knows no tribal bounds: the Old English epic *Beowulf* mentions the Gothic king Ermanaric, who lived in the Ukraine in the fourth century, while the Old High German *Hildebrandslied* evokes the fifth-century Italian wars of Odoacer and Theodoric. Thanks to the closeness of the languages and the richness of their attestation, scholars have been able to arrive at a remarkably detailed picture of Proto-Germanic, despite the absence of contemporary written sources. The technique that has made this possible is the comparative method, which has also been successfully applied to the reconstruction of Proto-Celtic, Proto-Italic, Proto-Greek, Proto-Indo-Iranian and Proto-Indo-European itself. It is probably safe to say that, despite some abiding problems, Germanic is one of the best understood branches of the IE family.

What kind of language, then, was Proto-Germanic? Like Celtic and Italic, its neighbors to the south and west – but unlike Balto-Slavic, its neighbor to the east – Germanic belongs to the *centum* division of IE languages; that is, it prehistorically merged the PIE "palatal" sounds **k̂*, **ĝ* and **ĝh* (phonetically perhaps /k^y/, /g^y/, /g^yh/) with the velars **k*, **g* and **gh*, keeping the resulting **k*, **g* and **gh* distinct from the laviovelars *$k^{u̯}$, *$g^{u̯}$ and *$g^{u̯}h$. In other respects, such as the merger of PIE **a* and **o* and the substitution of **m* for **bh* in the ending of the dative (< instrumental) plural, Germanic accords with Balto-Slavic. Phonetic and grammatical agreements of this kind are matched by lexical isoglosses : the Germanic words for « nourish » (OE *alan*) and « seek » (Go. *sokjan*) are shared only with Celtic and Italic (cf. OIr. *alid*, Lat. *alere*; OIr. *saigid* (« makes for »), Lat. *sāgīre*), while the name of the aspen tree (OHG *aspa*) and the word for « bare » (OHG *bar*) recur only in Balto-Slavic (cf. OPr. *abse*, Russ. *osina*; Lith. *bãsas*, OCS *bosŭ*). Such special correspondences, which can be multiplied, are normal between geographically contiguous

related languages and in no way compromise the position of Germanic as an autonomous branch of Indo-European.

2. Phonology

2.1. Consonants

On the phonological level, the most conspicuous development of the period between the breakup of Proto-Indo-European and the formation of Proto-Germanic was the *Germanic Consonant Shift*, often known, in honor of one of its co-discoverers, as *Grimm's Law*. This consisted of three distinct changes :

a) the PIE voiceless stops **p, *t, *k, *k^{u}*, when not preceded by an obstruent, became the voiceless fricatives **f, *þ, *x* (> later **h* in most environments), **xw* (> later **hw*) ;

b) the PIE voiced stops **b* (rare), **d, *g, *g^{u}* became the voiceless stops **p, *t, *k, *kw* ; and

c) the PIE voiced aspirated stops **bh, *dh, *gh, *$g^{u}h$* became the voiced fricatives **ƀ, *đ, *ǥ, *ǥw*, which often gave voiced stops in the attested languages.

Examples are very numerous : cf. Eng. ***foot*** (Go. ***fotus***), ***three*** (Go. ***þrija***), ***horn*** (Go. ***haurn***), ***what*** (Go. ***ƕiata***) beside Lat. ***pēs, trēs, cornu, quod*** ; Eng. ***tooth*** (Go. ***tunþus***), ***corn*** (Go. ***kaurn***), ***quick*** (Go. ***qius***) beside Lat. ***dēns, grānum, uīuus*** (< ***gwīwos***) ; Eng. ***bite*** (Go. ***beitan***), ***dare*** (Go. ***(ga)-daursan***), ***goose*** (NHG ***Gans***), ***warm*** (NHG ***warm***, with *w*- < **ǥw*-) beside Skt. ***bhid***- « split », ***dhr̥ṣ***- « be bold », ***háṃsa***- (< **ĝh*-) « (wild) goose », ***gharmá***- (< **$g^{u}h$*-) « hot drink » [3]. These changes were not restricted to word-initial position, but also took place word-internally. Here, however, the voiceless fricatives produced by Grimm's Law, together with the inherited voiceless sibilant **s*, were under certain conditions subject to the later rule known as *Verner's Law*. The effect of Verner's Law was to convert **f, *þ, *x, *xw* and **s* to the corresponding voiced fricatives **ƀ, *đ, *ǥ, *ǥw* and **z* when the preceding vowel did not bear the pre-Germanic (= PIE) movable accent. Thus, the Germanic word for « father », which

3. An aberrant (dialectal ?) development of **$g^{u}h$* to **ƀ* seems to underlie Eng. ***bid*** (Go. ***bidjan*** ; PIE **$g^{u}hedh$*- « ask ») and ***bane*** (OE ***bana***, OHG ***bano*** « slayer » ; PIE **$g^{u}hen$*- « slay »).

was originally end-stressed (cf. Skt. *pitár-*, Gk. *patḗr*), became **faþḗr* by Grimm's Law and **fađḗr* (> Go. *fadar*, OE *fæder*, OHG *fater*) by Verner's Law; the originally root-accented word for «brother», on the other hand, retained its **-þ-* < **-t-* until the end of the Proto-Germanic period (cf. Go. *broþar*, OE *brōðor*, OHG *bruoder*; Skt. *bhrā́tar-*, Gk. *phrā́tēr*). The conditioning environment for Verner's Law was lost when the pre-Germanic accentuation system, inherited from Proto-Indo-European, was replaced by the later Germanic system of fixed initial stress.

Grimm's and Verner's Laws are justly celebrated beyond the narrow circle of Germanicists and Indo-Europeanists. The discovery of Grimm's Law in the early nineteenth century furnished dramatic proof that sound change was a systematic process, thereby stimulating the search for other recurrent correspondence patterns and fostering the growth of the fledgling science of comparative philology. Later, in the 1870's, the discovery of Verner's Law helped to show that even the apparent exceptions to regular sound changes could be systematically explained – a recognition eventually enshrined in the famous principle that «sound laws have no exceptions». Yet despite many years of careful, and sometimes even obsessive, study, it is remarkable how little is known for certain about the mechanism or chronology of Grimm's and Verner's Laws. According to a widely held view, the initial step in the Germanic consonant shift was the passage of the voiceless stops to voiceless aspirates ($*p \rightarrow *p^h$, $*t \rightarrow t^h$, etc.), a change which left the voiced stops free to lose their voicing while the aspirated stops, both voiced and voiceless, developed further to fricatives. This is a plausible scenario, and it can be paralleled, at least typologically, by a similar shift in Armenian; but in the last analysis it is only a hypothesis. Other theories continue to command respect as well, including the view that the Germanic voiceless stops go back to glottalized **p', *t'*, etc., rather than to voiced **b, *d*, etc., in the parent language (the "glottalic theory"). As for the absolute chronology of these changes, all that can be said is that Grimm's Law took place before the bulk of the dialectal Germanic borrowings from Latin, but later than the acquisition of such loanwords as Go. *reiki*, OHG *rīhhi* «kingdom» (< Celtic **rīgio-*; cf. OIr. *ríge* «id.») and Go. *paida*, OHG *pfeita* «cloak» (Gk. *baítē* «shepherd's cloak», of foreign origin), which seem to have undergone the shift.

The effect of Grimm's and Verner's Laws was to leave early Proto-Germanic with an obstruent system that was far richer in fricatives than in stops. This imbalance, "unnatural" from a typological point of view, was partly remedied by subsequent changes, the most important being the development of the voiced fricatives to stops in certain environments. By the end of the Proto-Germanic period, **ƀ, *đ, *g, *gw* had become **b, *d, *g, *gw* after nasals, and **ƀ* and **đ* had probably become **b* and **d* in word-initial position as well. Go. *bindan* «tie» (cf. Skt. *bándhati* <

*bhándhati) was thus phonetically [bindan], although the contrast between [b], [d] and [ƀ], [đ] remained phonologically non-distinctive. Further fricative > stop rules were introduced in the individual Germanic dialects.

The remaining PIE consonants were the three "laryngeals" *h_1, *h_2, *h_3, and the six sonorants *r, *l, *m, *n, *i̯, *u̯, the latter with syllabic allophones *r̥, *l̥, *m̥, *n̥, *i, *u. As in most other IE languages, the laryngeals were normally lost in Germanic [4], although their earlier presence can be detected from such effects as compensatory lengthening and the "coloring" of adjacent vowels (cf. Go. *ietan* «let, release» < PGmc. *lētana^n < *lēd- < PIE *leh$_1$-d-; Go. *broþar* < PGmc. *brōþēr < *bhrā- < PIE *bhrah$_2$- < *bhreh$_2$-). Among the sonorants, the liquids (*r, *l) and the nasals (*m, *n) were preserved unchanged in most positions; new instances of consonantal *r, *l, *m, *n were introduced by the change of syllabic *r̥, *l̥, *m̥, *n̥ to *ur, *ul, *um, *un (cf. Go. *wulfs*, OE *wulf* « wolf » < *u̯l̥kʷo-/*u̯l̥po-; Go., OE *hund* « hundred » < *ḱm̥tóm). The treatment of the glides *i̯ and *u̯ was more complicated. *u̯ was stably maintained as PGmc. *w, but *i̯, which survived as *j in word-initial position and after consonants, was apparently lost after vowels other than *i (cf. Go. *salbon*, OHG *salbōn* « anoints », with *-ō- < *-āi̯e/o- < *eh$_2$-i̯e/o-). The much-discussed *Verschärfung* or "hardening" of intervocalic *-i̯- and *-u̯- to PGmc. *-jj- and *-ww- seems to have been phonologically regular when pre-Gmc. *-j- was preceded by *ī or an *i*-diphthong (i.e., *-ijja- < *-īja-, *-ajja- < *-aija-, etc.), or when pre-Gmc. *-w- was preceded by *ū or a *u*-diphthong (i.e., *-uwwa- < *-ūwa-, *-awwa- < *-auwa-, etc.). Such sequences frequently arose through the loss of a laryngeal, as in PGmc. *hawwana^n « hew, chop » (ON, *hǫggva*, OE *hēawan*) < *hauwa- < *hau'a- < *kauh$_2$-o-.

Proto-Germanic, then, had the following consonant system: [5]

voiceless stops	p	t		k
voiceless fricatives	f	þ	s	x/h
voiced obstruents	b/ƀ	d/đ	z	g/g̱
nasals	m	n		
liquids		r, l		
glides	w	j		

This inventory was preserved in Gothic, where the only important phonological innovations were the devoicing of word-final voiced fricatives (cf. nom. sg. *dags* « day » < *dagaz) and the change of *-ww-

4. A possible exception is to be seen in forms like OE *naca*, ON *nǫkkvi* "boat" (< PIE *neh$_2$u̯-ón-?) and Go. *unkis* "us two (dat. acc.)" (< *unkwis < PIE *n̥h$_3$-u̯é?), where *h_2 and *h_3 seem to have given PGmc. *k before *u̯ when the accent followed.

5. It is here assumed that *kw, *hw and *gw, the reflexes of the PIE labiovelars, patterned phonologically as clusters in Proto-Germanic.

and **-jj-* to *-ggw-* and *-ddj-* (cf. *triggwa* «faithful» < **triwwa-*, *twaddje* «two [gen. pl.]» < **twajjē*^{n}). In Old Norse the contrast between voiced and voiceless fricatives was lost altogether : **z* became *r* (cf. *dagr* «day»), while **f* and **þ* were voiced to [v] and [đ] after vowels and liquids, thus partly effacing the results of Verner's Law (cf. *bróðir* and *faðir*, both with *-ð-*). Old Norse also witnessed important changes to the glides : **h*, no longer synchronically interpreted as a voiceless fricative, was restricted to word-initial position (cf. *slá* «slay» vs. Go. *slahan*); single **w* and **j* were lost in a number of conspicuous environments (cf. *ormr* «worm», *ungr* «young» vs. Go. *waurms, juggs*); **-ww-* and **-jj-* became *-ggv-* and *-ggj-* (cf. *tryggvi* (= Go. *triggwa*), *tveggja* (= Go. *twaddje*)). Geminate consonants, already firmly established in Proto-Germanic (cf. **fulla-* «full» < **fulna-* < PIE **pl̥ h*^{1} *-nó-*; **wissa-* «known» < PIE **u̯it*^{s}*tó-* < **u̯id-tó-*), became much more common, especially in West Norse, through such rules as the assimilation of **-nþ, *-lþ-* to *-nn-, -ll-*, the assimilation of **-nt-, *-nk-* to *-tt-, -kk-*, and the change of **-kj-, *-gj-* to *-kkj-, -ggj-*.

Some of these developments were paralleled in West Germanic. Here, as in Old Norse, **z* was lost as an independent phoneme, disappearing in final position (cf. OS *dag* «day») and falling together with **r* elsewhere (cf. OS *wārun* «were» < **wēzun*). Word-internal **f* merged with **ƀ* in part of the West Germanic area, but **đ* became **d* everywhere, thus remaining distinct from **þ* even in languages like Old English, where the voiceless fricatives were subject to secondary voicing (cf. *brōðor* vs. *fæder*). An important West Germanic novelty was the gemination of all consonants except **r* before a following **j* – a development which led to the rise of characteristic morphophonemic alternations between single and double consonants (cf. OS *tellian* «say» < **taljan*; pret. *teliđa* < **taliđō*^{n}). The most conservative language of the group, at least from a phonological point of view, was Old Saxon. Old English and Old Frisian, among other innovations, palatalized the velar consonants to affricates before front vowels and **j* (cf. OE *licgan* [-ǰǰ-], OFr. *lidza* «lie» < WGmc. **liggjan*), reducing **g* before front vowels to the status of a palatal glide. In Old High German the West Germanic consonant system was utterly transformed by the *High German Consonant Shift*. This complex of changes, which began in the Upper German area in the sixth century, eventually spread in a less general form to the East, Rhenish and Middle Franconian dialects. The most characteristic feature of the shift was the passage of the voiceless stops **p, *t, *k* to affricates, which became long tense fricatives in postvocalic position. Examples are legion : cf. OHG *pfeffar* beside Eng. ***pepper***; OHG *zehan* (<z-> = a *ts*-like affricate), *wazzar* (←zz→ = a sibilant distinct from /s/) beside Eng. *ten*, *water*; OHG **chuninc** (<ch-> = [kx] (Upper German only)), ***mahhōn*** beside Eng. ***king***, ***make***. The voiced sounds were also affected : WGmc. **d* became **t* in Upper German and part of Middle

German (cf. _tag_ «day»), while WGmc. _*b/ƀ_, _*g/g_ and early OHG _*đ_ < _*þ_ tended to become _p, k_ and _t_ in Upper German.

The modern Germanic languages have unusually rich and complex vowel systems. The vowel system of Proto-Germanic, however, was notably simpler than most of its descendants. Proto-Indo-European had five short vowels – an inventory which Germanic reduced to four by merging PIE _*a_ and _*o_ as _*a_ (cf. Go. _akrs_ «field» beside Lat. _ager_, Gk. _agrós_ «id.»; Go. _ahtau_ «eight» beside Lat. _octō_, Gk. _oktṓ_)[6]. (PIE «_*ə_», in reality a subphonemic epenthetic vowel that developed in the vicinity of cluster-internal laryngeals, yielded _*a_ in initial and final syllables (cf. «father», ultimately < _*ph$_2$tḗr_) but disappeared elsewhere (cf. Go. _dauhtar_ «daughter» < _*dhuĝh$_2$tḗr_)). There was also confusion of the front vowels _*e_ and _*i_ in some environments. _*e_ was raised before a following _*i_ or _*j_ (cf. OHG _ist_ «is» < _*esti_), before tautosyllabic nasals (cf. OHG _hring_ «ring» < _*hringaz_ < _*hrengaz_ (> Finnish _rengas_ «id.»), and in non-initial syllabes (cf. OHG _(gi)nerit_ «saved (past ptcp.)» < _*nazidaz_ < _*nosetós_). These rules affected the inherited diphthongs as well. PIE _*ai, *oi_ and _*au, *ou_ fell together as PGmc. _*ai_ and _*au_ (cf. Go. _haihs_ «one-eyed» (: Lat. _caecus_ «blind»), _wait_ «knows» (: Gk. _(w)oîde_ «id.»), while PIE _*ei_ developed via _*ii_ to _*ī_ (cf. OHG _bīz(z)an_, Go. _beitan_ «bite» (<ei> = [ī]) < _*bheid-_). PIE _*eu_, the only remaining diphthong, was raised to _*iu_ before a following _i_-vowel (cf. OHG _biotan_ «offer» < _*beuđan_, 3 sg. _biutit_ < _*biuđiþ_).

Just as the short vowels _*a_ and _*o_ merged as PGmc. _*a_, the PIE long vowels _*ā_ and _*ō_ fell together as PGmc. _*ō_ (cf. Go. _broþar_ (<o> = [ō]) beside Lat. _frāter_, Gk. _phrā́tēr_; Go. _flodus_ «tide» beside Gk. _plōtós_ «floating»). There was no confusion, however, between PIE _*ē_ and _*ī_; _*ē_ consistently yielded a long open front vowel (phonetically probably [ǣ]), which was lowered further to _*ā_ in North and West Germanic (cf. Go. 3 pl. _setun_ (<e> = [ē]), OHG _sāzzun_ «sat»). The long diphthongs _*ēi, *ōi_, etc. were shortened before consonants in Germanic, as in most other IE languages ("Osthoff's Law"), but probably maintained their existence in absolute final position until the end of the Proto-Germanic period.

Although the loss of the distinction between _a_- and _o_-vowels reduced the stock of early Proto-Germanic vowels to four shorts and four longs, the complexity of the system was increased by subsequent changes. Toward the end of the Proto-Germanic period – or immediately thereafter, according

---

6. Here and below, "PIE _*a_" and "PIE _*o_" refer indifferently to "true" _*a_ and _*o_ and to the secondary _*a_ and _*o_ that resulted from the coloration of _*e_ by a neighboring laryngeal. Similarly, the "long vowels" are loosely understood to include etymological short vowels that were compensatorily lengthened when laryngeals disappeared in early pre-Germanic.

to some scholars – the number of short vowels was again brought to five by the lowering ("_a_-umlaut") of *_u_ to *_o_ before an *_a_ or *_ō_ in the following syllable. The effect of this rule can be seen, e.g., in the word for « horn », which appears in Old Norse, Old High German, Old Saxon and Old English as _horn_, even though the original preform was clearly *_hurna_^{n} < *_k̂r̥nó-_ [7]. *_i_ was occasionally lowered to *_e_ in the same environment, as in OHG, OE _nest_ "nest" < *_nizdó-_ (: Skt. _nīḍá-_, Lat. _nīdus_, etc.). A Proto-Germanic change which was purely phonetic, but which had phonological consequences in the daughter languages, was the development of the long nasalized vowels *_ā_^{n}, *_ī_^{n} and *_ū_^{n} from earlier *_an_, *_in_ and *_un_ before *_h_ (cf. Go. _þahts_ [-ā-], OHG _(gi)dāht_, OE _þōht_ « thought [past ptcp.] » < *_þanhta-_). Also important was the addition to the system of « *$\bar{e}_2$, » a closer vowel than the old *_ē_ < PIE *_ē_, from which it remained distinct in every Germanic language except Gothic. The origin of *$\bar{e}_2$ is one of the most controversial issues in Germanic phonology. Very few lexical items can be shown to have had this vowel in Proto-Germanic, and most of them were either borrowings, like Go. _Kreks_ « Greek » (cf. OE pl. _Crēcas_), or words of unclear etymology, like Go. _fera_, OHG _fēra, fiara_ « side ». A central role in discussions of the problem is played by the word for « here » (Go. _her_, ON _hér_, OHG _hēr, hiar_, OE _hēr_), the exact formation of which remains obscure [8].

Late Proto-Germanic thus had the following vowels in non-final syllables :

short	u		i	long	ū		ī	nasal	ū^n		ī^n
	o		e		ō		$\bar{e}_2$				
		a				ē				ā^n	
diphthongs	ai	au	eu	iu							

Gothic, eminently conservative in its treatment of the Germanic consonant system, was more innovative in its treatment of the vowels. The distinction between the short high vowels and short mid vowels was for all practical purposes lost : *_e_ and *_o_ became _i_ and _u_, which were then lowered ("broken") to [ɛ] and [ɔ], written <ai> and <au>, before _r_ and _h_ (cf. _wairþan_ « become » < *_wirþ-_ < *_werþ-_). Similarly, *_eu_ and *_iu_ fell together as _iu_, while the remaining diphthongs, *_ai_ and *_au_, were monophthongized to [ɛ̄] and [ɔ̄] – a change not indicated in Wulfila's orthography (_haitan_ "order" = [hɛ̄tan]). PGmc. *_ē_ and *$\bar{e}_2$ merged as [ē],

7. The corresponding Gothic form, _haurn_, is ambiguous, since PGmc. *_u_ would regularly have given _au_ ([ɔ]) in this position. It is possible that _a_-umlaut was a common North and West Germanic innovation not shared by Gothic.

8. The Neogrammarian derivation of PGmc. *_h_$\bar{e}_2$_r_ from PIE * _k̂ēi-r_ is almost certainly incorrect; see further p. 274.

written <e>. The nasalized vowels lost their nasality, establishing /ā/ as a new phoneme distinct from /ō/ and /ɔ̄/.

Old Norse maintained many of the phonological features that were lost in Gothic, though often with significant phonetic alterations. The five short vowels of Proto-Germanic were in principle retained; all, however, were subject to umlaut rules of various kinds (see below), and **e* was sometimes "broken" to *ja* or *jǫ*, as e.g., in *hjalpa* «help» (: OE *helpan*, Go. *hilpan*). The long vowels generally preserved their identity, with **ē* becoming *ā* (<á>), as everywhere outside Gothic. Of the diphthongs, only **au* survived as such; **ai* became *ei* (cf. *steinn* «stone» beside Go. *stains*), while **eu* and **iu* transfered their syllabicity to the second element, giving *jó* and *jú* (cf. *bjóða* beside Go. *biudan*). Massive reductions took place in unstressed syllables. Here length distinctions were obliterated, diphthongs were monophthongized, and all syllabic nuclei that resisted syncope were ultimately reduced to *a, i/e*, or *u/o*. Thus, e.g., the PGmc. 2 sg. form **wakaiz(i)* «you are awake» (= Go. *wakais*) yielded ON *vakir*, while the corresponding 1 sg. pret. **wakaiđō*^{n} «I was awake» gave ON *vakða*. New instances of **ā*^{n}, **ī*^{n} and **ū*^{n} were produced from short vowels before *-*ns*-. All nasalized vowels, old and new, eventually lost their nasal component, but **ī*^{n} and **ū*^{n} were first lowered to **ē*^{n} and **ō*^{n}, whence *é* and *ó* (cf. *ósk* «wish» < **wū*^{n}*skō* < **wunskō* < **u̯n̥̄skā* < **u̯n̥h₁sk-*)[9].

Of the West Germanic languages, Old High German was the most conservative. Except when subject to umlaut, the short vowels were fairly well preserved, even in unaccented syllables. Syncope, far less pervasive than in Old Norse, was sensitive to syllabic weight : pre-OHG **hōrita* «I heard» became *hōrta* (cf. ON *heyrða*), but **lagita* «I laid» remained as *legita* (contrast ON *lagða*). The most important changes were those affecting the long vowels and diphthongs. WGmc. **ē* (< *$ē_2$) and **ō* showed an early tendency to diphthongize to *ia, ie* and *ua, uo* (cf., e.g., *bruoder* beside Go. *broþar*), while a new *ē* and *ō* were generated from **ai* and **au* in unstressed and, under certain conditions, stressed syllables (cf. *lōn* «wage» beside Go., ON *laun*). Where they survived as diphthongs, **ai* and **au* became *ei* and *ou* (cf. *houbit* «head» beside Go. *haubiþ*); **eu* generally passed to *io*, remaining distinct from *iu* < **iu*. Most of these rules also operated in Old Saxon, except that here the diphthongization of **ē* and **ō* to *ia* and *uo* was confined to one of the two main dialects, and the change of **ai* and **au* to *ē* and *ō* was universal (cf. *stēn* «stone», *hōƀid* «head»). Unlike Old High German, Old Saxon sporadically developed nasalized vowels before the voiceless fricatives **s*, **f* and **þ* – an Ingvaeonic rule that applied with greater consistency in Old English (cf. OS *sāfto* «gently» < *sā*^{n}*fto* < **sanfto*). Both Old High German and Old Saxon merged the

9. It will be noted from this example that the "long" syllabic sonorants **r̥̄*, **l̥̄*, **m̥̄* and **n̥̄* fell together with their short counterparts in Germanic.

nasalized vowels with their oral counterparts before the beginning of the literary period.

In Old English the most important changes to the short vowels were the fronting of **a* to *æ* (cf. *dæʒ* "day" (<ʒ> = [j]) and the subsequent "breaking" of *æ, e* and *i* to the short diphthongs *ea, eo* and *io* before liquids and glides (cf. *beorg* « mountain » < **berg; slēan* « slay » < **sleahan* < **slæhan* < **slahan*)[10]. The long vowels, when not umlauted, were generally maintained, but WGmc. **ā* (< PGmc. **ē*) was fronted to *ǣ* (cf. *wǣron* « were » = OS *wārun*), and a new *ā* was produced by the monophthongization of **ai* (cf. *stān* « stone »). The *u*-diphthongs remained diphthongal and contrastively long : **au* developed via **æw* and **eaw* to *ēa* (cf. *hēafod* « head »), while **eu* and **iu* gave *ēo* and *īo*, which later fell together in the dominant West Saxon dialect. As in Old Norse, there were far-reaching shortenings, reductions and losses in non-initial syllables, the effect of which was to leave *a, u/o* and *e* [ə] as the only completely unstressed vowels in the language. Nasalized vowels were generated before **-ns-, *-nþ-* and **-nf-* as well as before **-nh-; *ā*^{n} was rounded to **ō*^{n} and became *ō* with the loss of distinctive nasalization (cf. *sōfte* beside OS *sāfto*). The treatment of the vowels in Old Frisian was broadly similar.

All the Germanic languages other than Gothic show the effects of umlaut rules, which caused a vowel, normally stressed, to assimilate in one or more features to a vowel or consonant that followed. A phonetic tendency toward such assimilation was probably already a feature of Proto-Germanic; early illustrations can be seen in the raising of **e* to **i* before **i* and **j* and the lowering of **u* to **o* before **a* and **ō*. By far the most conspicuous type of umlaut in the daughter languages, however, was *i-umlaut* proper, which converted *a*-like vowels to *e*-like vowels, and back rounded (*u*-like and *o*-like) vowels to front rounded (*ü*-like and *ö*-like) vowels before **i, *ī* or **j* in the next syllable. There is no trace of this rule in Gothic or early Primitive Norse, but its workings can be seen in every later Germanic language. In general, *i*-umlaut was slower to establish itself in the continental West Germanic dialects than elsewhere. In early Old High German the process was still purely phonetic: vowels were fronted before a preserved *i* or *j*, but not before an etymological **i* which was lost by sound change (cf. *gast* « guest » < **gastiz*, but nom. pl. *gesti* < **gastīz*). In most cases the effects of *i*-umlaut were not shown in Old High German spelling, but began to be graphically represented in Middle High German, where the weakening of unstressed *i* to *e* [ə] rendered the new umlaut vowels contrastive (cf. *scōni* « bright, fair » > MHG *schœne* > NHG *schön*). By the middle of the eleventh century all the Old High German back vowels and diphthongs had acquired phonologically distinctive fronted counterparts, as

10. The precise environments for this rule were complex and varied considerably from dialect to dialect.

was also the case in Old Saxon/Low German. In Old Low Franconian and Dutch *i*-umlaut remained inconspicuous and purely phonetic, eventually disappearing with the loss of *i* in unaccented syllables.

In Old English and Old Norse the phonologization of *i*-umlaut was very early. The word for «guest» in these languages shows umlaut in the singular as well as the plural (OE nom. sg. *ʒiest*, ON *gestr* < *-*iz*; cf. Prim. Norse -*gastiR*); umlaut is also found in preterites of the type OE *hīerde*, ON *heyrða* «heard» < **hauziđō*^{n}, where the -*i*- that triggered the rule was syncopated after heavy syllables. Old Norse generally maintained its newly acquired rounded vowels and diphthongs (*y, ý, ø, ǿ, ey*), but Old English (West Saxon) kept only *y* and *ý*, unrounding the umlaut products of *ō* and *īo* to *ē* and *īe* (cf. *fēt* (dialectal *fǣt*, ON *fǿtr*) "feet" < **fōtiz*). A different but related process in Old Norse was *labial umlaut* (often separated into "*u*-umlaut" and "*w*-umlaut"), which induced rounding effects in stressed vowels similar to the fronting effects of *i*-umlaut (cf. nom. pl. *lǫnd* « lands » < **landu* < **landō*; *søkkva* «sink (tr.)» < **sekkwa* < **sankwijan* (= Go. *sagqjan*, OHG *senken*)). The term "*u*-umlaut" is also applied to a special breaking process in Old English (cf. *siolfor, seolfor* «silver» < **siluƀra-*).

The far-reaching changes of the immediate post-Germanic period left the medieval Germanic languages with exceptionally complex, "crowded" vowel systems. It is no accident that in the later development of Germanic these configurations were often unstable. "Chain shifts", marked by fresh diphthongizations, monophthongizations, frontings and/or raisings, dot the history of the modern languages. The best known example is the so-called Great Vowel Shift in English, which began in the fifteenth century, with the diphthongization of *ī* and *ū* to [ei] (later [ai]) and [ou] (later [au]), and the raising of the Middle English high mid vowels *ẹ̄* and *ọ̄* to fill the resulting structural gap. No other language shows precisely the same sequence of changes, but Standard German implemented a similar set of rules, diphthongizing MHG *ī, ū* and *ǖ* to [ei], [ou] and [öü] (later [ai], [au], [ɔi]), and then producing new instances of *ī, ū* and *ǖ* from the old diphthongs *ie, uo* and *üe*. An entirely different chain shift, triggered by the fronting of *u* to [ü], took place in Swedish and Norwegian.

The treatment of final syllables raises especially difficult problems in Germanic. Germanic was one of the branches of Indo-European which, like Balto-Slavic, Greek and Indo-Iranian, distinguished between two types of long final syllables. "Acute" finals, better termed "bimoric", were normally subject to shortening; thus, the bimoric accusative singular of *ō*-stem nouns, which originally ended in *-*ō*^{n} (< *-*ōn* < *-*ām* < PIE *-*eh*$_2$*m*), gave forms in (short) -*a* in Gothic (cf. *giba* «gift (acc.)») and -*a* in Old High German (*geba* «id.»), corresponding to Greek -*ān* (-*ā́n* under the accent) and Sanskrit -*ām*. Contrasting with these were "circumflex" or "trimoric" syllables, which maintained their length with greater tenacity.

Such a trimoric ending was the _*-ō̃_^{n} – note the use of the circumflex accent – of the Proto-Germanic genitive plural (cf. Gk. _-ōn_ (_-ō̃n_), Skt. _-ām, -aam_), which survived as (long) _-o_ in Gothic (_gibo_) and as _-o_ in Old High German (_gebono_). Trimoric finals arose in late Proto-Indo-European and early Proto-Germanic from earlier sequences containing a hiatus, typically the result of a lost laryngeal. The introduction of the bimoric : trimoric opposition in final syllables marked a complication in the Germanic phonological system _vis-à-vis_ that of the parent language, but the number of possible contrasts in final syllables was reduced in other ways. All final obstruents other than _*s_ and _*z_ were lost, and final nasals disappeared with nasalization of the preceding vowel. This left _*s, *z, *r, *l_ and _*n_ (< _*-nt_) as the only consonants permitted to stand at the end of a Proto-Germanic word.

Final syllables were lost or reduced in every Germanic language, but the chronological details are unclear. The apocope of short vowels other than _*u_ in absolute final position seems to have had its beginnings in late Proto-Germanic (cf. Go. _wait_ « knows », ON _veit_, OHG _weiz_, OS _wēt_, OE _wāt_ < _*wait(i)_ < _*u̯óide_). On the other hand, the syllable types _*-V̆_^{n} and _*-V̆C_ were retained into the dialectal period, as can be seen from forms like Prim. Norse _horna_ « horn » < _*horna_^{n} and _-gastiR_ < _*gastiz_. Gothic instituted a _Dreimorengesetz_ (« Law of Three Moras »), by which most final syllables were abbreviated by one mora : short vowels disappeared, bimoric long vowels were shortened, and trimoric long vowels were reduced to ordinary longs (cf. _dags_ « day » < _*-az_, _giba_ « gift » < _*-ō_, _galeiko_ « likewise » < _*-ō̃_). North and West Germanic weakened (bimoric) _*-ō_ to _*-u_ soon after their separation from Gothic ; this vowel, like all short-vowel endings that survived the first wave of truncations, was eliminated or retained according to circumstances in the individual dialects (cf. ON _gjǫf_ vs. OE _ʒiefu_ < _*geƀu_ < _*-ō_, but after a heavy syllable ON _rūn_ « rune, secret », OE _rūn_ < _*rūnu_ < _*-ō_). The trimoric endings were eventually reduced as well, but later ; no North or West Germanic language completely elided the _*-ō̃_^{n} of the genitive plural (ON _-a_, OHG _-o_, OS _-o_, OE _-a_) or the _*-ō̃_ of adverbs like Go. _galeiko_ (ON _-a_, OHG _-o_, OS _-o_). In every language except Icelandic and Faroese the tendency of vowels in final syllables to weaken or disappear persisted into the modern period, and went hand in hand with the decline of inflectional morphology.

3. Morphology and Syntax

3.1. Declension

In comparison with the turbulent history of Germanic phonology, the historical morphology of Germanic seems almost prosaic. Particularly in

its nominal system, Proto-Germanic was a fairly typical IE dialect, not unlike, e.g., Latin, its approximate contemporary. There were three genders (masculine, feminine and neuter) and two numbers (singular and plural), the dual having been lost except in the personal pronouns. Of the eight PIE cases, Germanic preserved the nominative, accusative, vocative, genitive and instrumental; the locative, perhaps with some admixture from the dative, gave the synchronic "dative", while the ablative survived only in adverbs. No individual language, however, actually retained six cases. The vocative, proper to the singular alone, disappeared everywhere except in Gothic; the instrumental survived as a case-form distinct from the dative only in a few noun classes in West Germanic, and here again only in the singular. Noun paradigms in the older Germanic languages typically consist of a nominative, accusative, genitive and dative, sometimes with the addition of a vocative or instrumental in the singular.

The inflection of PIE nominal stems was a complex process, commonly involving accent movement and stem-internal ablaut (apophony) as well as the addition of case endings (cf. e.g., nom. sg. **mén-ti-s* (later **mn̥-tí-s*) «mind», gen. **mn̥-téi-s*, loc. **mn̥-tḗi*). Most of this apparatus was eliminated in Proto-Germanic. Accentual mobility was lost by sound law; all words acquired fixed initial stress. The role of ablaut was drastically reduced, partly through the expansion of the apophonically invariant *a*-stems and *ō*-stems (= PIE *o*-stems and *ā*-stems), and partly through the synchronic reanalysis of inherited forms with suffix ablaut (cf. Go. dat. sg. *gamundai* «mind», historically < **mn̥-tḗi-Ø* but synchronically *-mun-d-ai*). Descriptively speaking, Proto-Germanic nouns fell into five major and several minor declensions, each characterized by a distinctive set of synchronic endings and conventionally named for the vowel or consonant that historically ended the stem (*a*-stems, *ō*-stems, *i*-stems, *u*-stems, *n*-stems). Following a venerable tradition initiated by Grimm, the vocalic stems are sometimes called "strong", while the *n*-stems are called "weak".

A good impression of noun declension in the older Germanic languages is afforded by the Gothic, Old Norse and Old High German paradigms of the masculine *a*-stem **dag̱az* «day» :

	Go.		**ON**		**OHG**	
	sg.	pl.	sg.	pl.	sg.	pl.
nom.	*dags*	*dagos*	*dagr*	*dagar*	*tag*	*tagā̆*
gen.	*dagis*	*dage*	*dags*	*daga*	*tages*	*tago*
dat.	*daga*	*dagam*	*degi*	*dǫgum*	*tage*	*tagum*
acc.	*dag*	*dagans*	*dag*	*daga*	*tag*	*tagā̆*
voc.	*dag*	(= nom.)	(= nom.)		(= nom.)	
instr.	(= dat.)		(= dat.)		*tagu*	(= dat.)

These forms, corresponding etymologically to the Latin and Greek second (*o*-stem) declension, can in most cases be easily traced to Proto-

Germanic and PIE antecedents. In the singular, the nominative, accusative, vocative and instrumental go directly back to PGmc. **-az* (PIE **-os*), **-a*^{n} (PIE **-om*), **-i* (PIE **-e*) and **-ō* (PIE **-ō* < **-oh*$_1$), respectively. The gen. sg. in Old High German and Old Norse continues PGmc. **-asa* < PIE **-oso*, a variant of **-osi̯o* (= Gk. *-oio*, Skt. *-asya*); Go. *-is* is a borrowing from the pronominal declension, where, e.g., the gen. sg. *þis* « of this, of that, *eius* » continues PIE **teso*. The dat. sg. in Gothic probably represents the old instrumental; in Old High German and Old Norse (where *degi* underwent the rule of "*g*-umlaut") the corresponding ending is best taken from the PGmc. loc. sg. in **-aĩ* (trimoric ?) < PIE **-oi*̆. Of the plural forms, the Gothic and Old Norse nominative and accusative go back to **-ōz* (or **-ō̃z*; PIE **-ōs* < **-o-es*) and **-ans* (PIE **-ons*), one or both of which presumably also underlies the somewhat mysterious OHG *-ā̆* . The gen. pl. in Old Norse and Old High German continues PGmc. **ō̃*^{n} (PIE **-oHom* ?), replaced under obscure circumstances by **-ē̃*^{n} in pre-Gothic. The dat. pl. forms are, as always in Germanic, instrumentals in **-mi(z)* < dialectal PIE **-mi(s)*, with **-m-* for **-bh-* as in Balto-Slavic [11].

There were also neuter *a*-stems, which maintained the Indo-European peculiarity, common to all neuters, of making no distinction between nominative and accusative (cf. Go. nom.-acc. sg. *land* « land » < **-a*^{n} < **-om*, nom.-acc. pl. *landa* < **-ō* < **-ā*; ON *land*, pl. *lǫnd* < **-u*; OHG *lant*, pl. *lant* < **-u*). A few neuter *a*-stems in West Germanic had plurals in **-ir(u)* (cf. OE *cild* "child", pl. *cildru*; OHG *lamb* « lamb », pl. *lembir* (= NHG *Lämmer*)); these were originally *s*-stems (cf. Lat. *genus* « race », pl. *genera* (< **-esa*)) in which the formative element *-ir-* < **-iz-* < **-es-* was synchronically reinterpreted as a number marker. Very important from a descriptive point of view were the *ja*-stems, both masculine and neuter, which continued PIE stems in **-i̯o-* and **-ii̯o-*. In Proto-Germanic the distribution of **-ja-* (< **-i̯o-*) and **-ija-* (< **-ii̯o-*) was metrically regulated: light syllables were followed by **-ja-*, while heavy syllables and polysyllabic stems were followed by **-ija-*. The difference between the two suffix-variants, ultimately a reflex of the PIE rule known as Sievers' Law, appears clearly in Go. *harjis* « army » (< **kori̯o-*) vs. *hairdeis* « shepherd » (< **k̂erdhii̯o-*). *ja*-stem inflection was productive in Germanic, being associated, e.g., with the borrowed suffix **k̂ā̆rija-* < Lat. *-ārius*, which was later exploited as a device for forming agent nouns (cf. Go. *bokareis* « scribe », OE *bōcere* « id. », OHG *lērāri* (= NHG *Lehrer*) « teacher »).

The other strong declensions present few problems. The *ō*-stems, like their PIE counterparts, the *ā*-stems, were exclusively feminine. Many were abstracts, including the common deadjectival nouns in **-iþō* (cf. Go. *mildiþa* « mercy », OHG *miltida* beside OHG *milti* « mild, merciful »). A

11. The PIE instr. pl. of *o*-stems ended in **-ōis*; PGmc. **-amiz* was created on the model of the other stem-classes.

distinct subclass were the *jō*-stems, some of which preserved an archaic nom. sg. in PGmc. *-ī*, recalling the Sanskrit type *devī* «goddess», gen. *devyās* (cf. Go. *bandi* «bond», gen. *bandjos*). The *i*-declension originally contained both masculines and feminines, which had identical paradigms in the parent language. Over time, however, the two diverged; in every Germanic language the masculine *i*-stems were analogically influenced by the *a*-stems, while the feminines, which included the common deverbal abstracts in *-điz* (*-tiz*), tended to keep their inherited endings (cf. Go. *gasts*, gen. *gastis* [m.] vs. *gahugds* «thought», gen. *gahugdais* [f.]). Both genders retained *-i-* or *-j-* in the endings of the plural, a fact ultimately responsible for the rise of the umlaut plural types *Gast : Gäste* and *Kraft : Kräfte* (< OHG *kraft* [f.] «strength», pl. *krefti*) in Modern German. The *u*-stems, containing words of all three genders, were at the outset a relatively small class. In Gothic, however, the inherited stock of *u*-stems was considerably augmented by loanwords (e.g., *aggilus* «angel») and former root nouns, which developed an *-u-* by sound change in the acc. sg. (cf. *fotus* «root», acc. *fotu* < *-u^n < *-*un* < *-*m̥*).

Most of the early Germanic dialects preserve traces of consonant-stem inflection in the kinship terms (*r*-stems; cf. Go. *broþar*, gen. *broþrs*, dat. *broþr*), in petrified present participles (*nt*-stems; cf. Go. *fijands* «enemy», dat. *fijand*, nom. pl. *fijands*) and in scattered root nouns (cf. Go. *baurgs* «city», gen. *baurgs*, dat. *baurg*). None of these groups, however, compares in importance with the *n*-stems, the expansion of which was one of the most characteristic morphological innovations of Germanic. The masculine *n*-stems were an inherited class, originally marked by an ablauting suffix *-*on*-/*-*en*-/*-*n*-, with a nom. sg. in PIE *-*ō* < pre-PIE *-ōn* (cf. Lat. *homo*, *-inis* «man»). (A less numerous subtype had only *-*en*- and *-*n*-, with *-*ēn* in the nom. sg. (: Lat. *pecten*, *-inis* «comb»)). In pre-Germanic the nom. sg. in *-*ō* was remade to *-*ōn*, which became PGmc. *-$\bar{o}^n$; the other cases added the consonant-stem endings, which were mostly of the form *-$\breve{V}$, *-$\breve{V}^n$ or *-$\breve{V}C$, to the *-*n*- of the stem. The result was a paradigm in the daughter languages in which the nom. sg. ended in a vowel and the oblique cases tended to end in simple *-n*, with no desinence proper (cf. Go. *guma* "man", gen. *-ins*, dat. *-in*, acc. *-an*, nom.-acc. pl. *-ans*; OHG *gumo, -en, -en, -on, -on*; OE *guma, -an, -an, -an, -an*). Each language had its own peculiarities: the West Germanic dialects collectively substituted trimoric *-$\tilde{\bar{o}}^n$ for -$\bar{o}^n$ in the nom. sg., probably under the influence of the weak adjectives (see below); Old English generalized the *o*-grade suffix-form *-an-* to all the oblique cases; Old Norse replaced *-$\bar{o}^n$ by *-$\bar{e}^n$ in the nom. sg. and lost *-*n* by sound change elsewhere (*gumi, -a, -a, -a*, pl. *-ar*). The neuter *n*-stems were likewise an old type, the inherited membership of which included the PIE word for «name» (cf. Go. *namo* < *-$\tilde{\bar{o}}^n$, gen. *-ins*; cf. Lat. *nōmen, -inis*). Wholly new in Germanic, on the other hand, were

the feminine *n*-stems, which were of two types : a) *n*-extended forms of original *ō*-stems (< *ā*-stems), such as Go. *tuggo* « tongue », gen. *tuggons*, dat.-acc. *tuggon* (= OHG *zunga, -ūn*; OE *tunge, -an*; cf. OLat. *dingua*, Skt. *jihvā́*); and b) similarly extended forms of original feminines in *-*ī*-, which served as adjectival abstracts (cf. Go. *hauhei* « height », gen. *-eins*; OHG *hōhī*, gen. *hōhī(n)*). In the later West Germanic languages the *n*-stems were the starting point for the common plural types in *-en* (cf. Eng. *oxen*, NHG *Zungen*, etc.).

The declension of pronouns in Proto-Germanic, as in every other IE language, was full of anomalies. The personal pronouns were highly irregular, but retained such archaisms as separate dual forms in the first and second persons (cf. p. 257, fn. 4) and a suppletive paradigm based on the PIE stems **u̯ei-* and **nes-* in the 1 pl. (cf. Go. *weis* « we » vs. *uns* « us », recalling Skt. *váyam : nas*). The possessive adjectives **mīna-* « my », **þīna-* « thy », **sīna-* « his, her, their (own) », **unsara-* « our », **izwara-* « your » were distinctively Germanic creations. Among the demonstratives, Germanic retained the PIE **so-/*to-* pronoun (cf. Go. nom. sg. masc. *sa* « this, that », fem. *so*, nt. *þata*), which became a definite article in West Germanic, as well as the pronoun **e-/ *i-*, which furnished personal pronouns, especially in Gothic and Old High German (cf. Go. *is, ita* « he, it », OHG *er, iz*). Most Germanic languages also made significant use of the PIE demonstratives **k̂e-/*k̂i-* (cf. OE, OS *hē̆* « he », ON *hann* « id. » and **eno-*, which became a definite article in Old Norse (nom. sg. masc. *-inn*). All these shared a number of inflectional peculiarities, such as the presence of a *t*-element in the nom.-acc. sg. neuter (Go. *þata*, OHG *daz*; cf. Skt. *tád*), an ending **-nō*ⁿ in the acc. sg. masc. (Go. *þana*, OHG *then*), a stem-enlargement **-mm-* < **-zm-* in the dat. sg. masc. and nt. (Go. *þamma*, OHG *themu*; cf. Skt. *tásmai*), and an enlargement **-z-* in the gen. and dat. sg. fem. and gen. pl. of all genders (Go. *þizos, þizai, þize/þizo*; cf. Skt. *tásyās, tásyai, téṣām/ tā́sām*)). The interrogative stem **hwa-* (Go. *ƕas* « who », OHG *(h)wer*, etc.), though lacking a plural, displayed many of the same features. There was no separate relative pronoun in Germanic, the PIE relative **i̯o-* (cf. Gk. *hós*, Skt. *yá-*) having disappeared.

Adjectives were provided with a dual inflection in Proto-Germanic. Depending on their immediate syntactic environment, they either remained "strong", being declined as *a/ō*-stems, *ja/jō*-stems, *i*-stems or (occasionally) *u*-stems; or they became "weak", acquiring the declension of *n*-stems. The weak declension of adjectives arose from the use of **-on-/*-en-/*-n-* as an "individualizing" suffix in late Proto-Indo-European – a function seen in Greek and Latin personal names of the type Gk. *Strábōn*, gen. *Strábōnos* « Squint-eyes » < *strabós* « squint-eyed » and Lat. *Cato, -ōnis* « Smarty » < *catus* « smart, sharp ». PGmc. **gōđaz* thus originally meant simply « good [nom. sg. masc.] », while **gōđō̃ ⁿ* (i.e., **gōđa-* + *-ō*ⁿ) meant

« the good one ». The distinctive innovation of Germanic was to reinterpret the semantic difference between the two in purely syntactic terms, so that **gōđō̃^{n}* became obligatory in the inherently individualizing environment of a preceding demonstrative, even when a noun followed (**sa gōđō̃^{n} gumō^n* « the good man »). In this way every adjective acquired a complete set of *n*-stem forms in all three genders. Predicate adjectives and adjectives in indefinite noun phrases were unaffected by the new rule and continued to inflect as vocalic stems (**gōđaz gumō^n* « a good man »), though their individual declensional forms were in many cases borrowed from the pronouns (cf. Go. dat. sg. masc. *godamma*, acc. sg. *godana*, gen. sg. fem. *godaizos*, etc.; cf. NHG *gutem, guten, guter*)[12] It is of interest to note that a typologically similar distinction between "definite" and "indefinite" adjectives was created, though with very different formal means, in Balto-Slavic.

3.2. Conjugation

3.2.1. Strong verbs

The history of the Germanic verbal system is for the most part a history of simplification and regularization. Proto-Indo-European had a present and three preterite-like tenses : the imperfect, built from the present stem and thus marked for durativity or iterativity (cf. Gk. *éthnēiskon* « I was dying »; the aorist, properly denoting a punctual, and hence non-present, action (*éthanon* « I died »); and the perfect, originally denoting a state but frequently taking on the value of a resultative preterite (*téthnēka* « I am dead » > post-class. « I have died »). Germanic merged all three into a single category, known simply as the preterite. The preterite and the present, which served also as a future, were the only tenses in the Proto-Germanic verbal system; there were no periphrastic formations of the type found in the medieval and modern Germanic languages (cf. Eng. *I will do, I have done*, etc.). In its treatment of the inherited categories of mood, voice, person and number, Germanic was an averagely conservative branch of the IE family. Of the moods, the subjunctive was lost, but the indicative, imperative and optative were retained. The PIE mediopassive was preserved as a grammatical passive, at least in the present (cf. Go. *bairada* < **bhérotoi*; cf. Gk. *phéretai*); a trace of the older middle value survives in Go. 1, 3 sg. *haitada* « am, is called », OE *hātte* « id. », Prim.

12. Comparative adjectives, made with the zero-grade of the PIE suffix **-i̯os-/ *-i̯es-/ *-is-*, lacked strong forms entirely (cf. Go. *batiza* « better », ON *betri*, OHG *bezziro*). Superlatives had the usual twofold declension.

Norse 1 sg. *haite* (< PGmc. **haitō̃i*) « am called ». Corresponding to the special dual pronouns of the first and second persons, Proto-Germanic retained distinct first and second person dual forms of the verb. In the historical period, however, the passive and the dual were confined to Gothic. The earliest North and West Germanic languages expressed the passive by syntactic means and used plural verb forms with dual subjects.

Presents were formed in a variety of ways in the parent language. Some verbal roots, for example, made presents without any kind of affixation (e.g., **h₁ḗd-ti* [root **h₁ed-*] « eats »); others made reduplicated presents (e.g., **dhédheh₁-ti* [root **dheh₁-*] « puts » or nasal-infix presents (e.g. **bhi-né-d-ti* [root **bheid-*] « splits »); still others made "thematic" presents in **-e/o-* (e.g., **bhér-e-ti* [root **bher-*] « carries ») or **-i̯e/o-* (e.g., **bhudh-i̯é-tor* [root **bheudh-*] « is alert, notices »). Germanic, like most of its sister branches, greatly extended the range of the thematic formations; in particular, it regularized and generalized the simple thematic type with *e*-grade of the root (**bhéreti*). The overwhelming majority of primary verbs were provided with such presents, which often replaced older formations. Thus, e.g., **h₁éd-ti* was in effect supplanted by **édeti* (Go. *itiþ* « eats »); **bhi-né-d-ti* became **bhéideti* (Go. *beitiþ* « bites »); **bhudh-i̯é-tor* became **bhéudheti* (Go. *-biudiþ* « offers, orders » < « remarks »). Thematic presents were also built to roots which had no synchronic *e*-grade, such as PGmc. **an-* « breathe » (cf. Go. *-aniþ* < **áneti* < **h₂enh₁-* or PGmc. **skaƀ-* « scrape » (cf. Go. *skabiþ* < **skábheti* < **skabh-*). Remaining vestiges of the non-thematic present classes were generally thematized, as, e.g., Go. *standiþ* « stands », with both nasal infix (cf. OE pret. *stōd*, Eng. *stood*) and stem-final **-e/o-*. The only competing present formation of any prominence was the type in **-i̯e/o-*, which was especially common in verbs with Germanic *a*-vocalism (cf. Go. *hafjiþ*, OHG *hevit* « lifts » (= Lat. *capit* < **káp-i̯e-ti* « takes »).

The inflection of simple thematic presents in the Germanic dialects can be illustrated by the following forms of the verb « to become » (root PGmc. **werþ-* < PIE **u̯ert-* « turn ») :

			Go.	ON	OHG	OE
indic.	sg.	1	*wairþa*	*verð*	*wirdu*	*weorþe*
		2	*wairþis*	*verðr*	*wirdis*	*wierst*
		3	*wairþiþ*	*verðr*	*wirdit*	*wierþ*
	pl.	1	*wairþam*	*verðum*	*werdumēs*	*weorþaþ*
		2	*wairþiþ*	*verðið*	*werdet*	*weorþaþ*
		3	*wairþand*	*verða*	*werdant*	*weorþaþ*
impv.	sg.	2	*wairþ*	*verð*	*wird*	*weorþ*
	pl.	2	*wairþiþ*	*verðið*	*werdet*	*weorþaþ*
opt.	sg.	2	*wairþais*	*verðir*	*werdēs*	*weorþe*

	3	*wairþai*	*verði*	*werde*	*weorþe*
pl.	1	*wairþaima*	*verðim*	*werdēm*	*weorþen*
	3	*wairþaina*	*verði*	*werdēn*	*weorþen*
ptcp.		*wairþands*	*verðandi*	*werdanti*	*weorþende*
infin.		*wairþan*	*verða*	*werdan*	*weorþan*

The endings of the Proto-Germanic indicative forms were best preserved in Gothic. Gothic retained the original distribution of *e*- and *o*-color in the thematic vowel, with *-i-* < **-e-* in the 2 sg. (PIE **-esi*), 3 sg. (**-eti*) and 2 pl. (**-ete*), and *-a-* < **-o-* elsewhere (1 sg. **-ō* < **-o-*h_2, 1 pl. **-o-me*, 3 pl. **-o-nti*). The 2 sg. in *-is* and 3 sg., 2 pl. in *-iþ* go back to **-iz* and **-iđ*, with voicing by Verner's Law, as shown by the preserved *-d* of the 3 pl. in *-and*. Voiced endings were also generalized in Old Norse, as is clear from the 2 sg. in *-r* < **-iz*; a striking Old Norse innovation was the substitution of *-r* for **-ð* in the 3 sg., a change datable to the seventh century. Old High German preserved a Proto-Germanic feature not found in Gothic or Old Norse – the raising of the root vowel *-e-* to *-i-* in the 2 sg. and 3 sg., where the following ending contained an *-i-*. (A later extension of the same process brought about the change of **werþu* to *wirdu* in the 1 sg.) The OHG 1 pl. in *-umēs* is noteworthy : the development of *-u-* from *-a-* in medial syllables before *-m-* was phonologically regular in North and West Germanic (cf. ON *-um*), but the sequence *-mēs* remains unexplained. Old English 2 sg. *wierst* and 3 sg. *wierþ* presuppose earlier **wirþis* and **wirþiþ*, with the voiceless variants of the endings (cf. OS *-is, -iđ*); the addition of *-t* in the 2 sg., a feature later found in other West Germanic dialects as well (cf. MHG *-st*), was due in part to the influence of the 2 sg. pronoun. An Ingvaeonic trait, common to Old English and Old Saxon, was the generalization of the 3 pl. in **-anþ* (> OE *-aþ*, OS *-ađ*) to the other plural persons.

Of the non-indicative forms, the 2 sg. imperative goes back to the bare thematic stem in **-i* < **-e*, as in most early IE languages[13]. The optative forms point to Germanic prototypes in **-ai-* (< PIE **-o-ih*$_1$*-*), which survived as (graphic) *-ai-* in Gothic and underwent monophthongization and/or weakening elsewhere. The endings of the optative were partly distinct from those of the indicative, reflecting the fact that the optative in the parent language was associated with the "secondary" endings, which lacked final **-i*. The Gothic 3 sg. in *-ai* thus corresponds exactly to Skt. *-et* (*bháret*) and Gk. *-oi* (*phéroi*), while the 3 pl. in *-aina* shows the secondary addition of a final vowel, presumably taken from the 1 pl., to an original ending **-a(j)in* < **-o̯ient* (cf. Gk. *phéroien*, Avestan *baraiiən*). The Ingvaeonic languages generalized the ending of the 3 pl. to

13. Gothic has a 3 sg. impv. in *-adau* and a 3 pl. impv. in *-andau*, the exact relationship of which to the corresponding forms in Sanskrit and Hittite remains to be clarified.

the other persons, as in the indicative; Old High German kept the plural forms distinct, but with *-m*, not *-mēs*, in the 1 pl. All the Germanic languages inherited a present participle in **-and-* < **-ont-* (cf. Gk. *phéront-*), which was enlarged to an *n*-stem in Gothic and Old Norse (Go. nom. sg. *wairþands* or *-anda*, gen. *-andins*, dat. *-andin*, etc.), but to a *ja*-stem in West Germanic. The attested infinitives in *-an* (ON *-a* < **-an*) go back to a PGmc. verbal noun in **-ana*^{n} < **-onom*, comparable to the Sanskrit verbal abstracts in *-anam* (*e.g., bháraṇam* «carrying»).

The preterites of the primary verbs in **-ana*^{n} and **-jana*^{n} – or "strong" verbs, as they are called – were based on the PIE perfect. Originally a stative category of restricted distribution, this formation became productive in Germanic, encroaching upon and eventually usurping the functions of the imperfect and aorist. The perfect in the parent language was marked by three distinctive characteristics: reduplication, a special set of personal endings, and a distinctive ablaut pattern that opposed an accented *o*-grade root in the singular to an unaccented zero-grade root in the plural. Of these three features, reduplication was usually lost in Germanic, but the ablaut, accent and endings of the PIE perfect were retained. The result was the Germanic strong preterite, of which the verb **werþana*^{n} again furnishes a typical example:

			Go.	ON	OHG	OE
indic.	sg.	1	*warþ*	*varð*	*ward*	*wearþ*
		2	*warst*	*vart*	*wurti*	*wurde*
		3	*warþ*	*varð*	*ward*	*wearþ*
	pl.	1	*waurþum*	*urðum*	*wurtum*	*wurdon*
		2	*waurþuþ*	*urðuð*	*wurtut*	*wurdon*
		3	*waurþun*	*urðu*	*wurtun*	*wurdon*
opt.	sg.	2	*waurþeis*	*yrðir*	*wurtīs*	*wurde*
		3	*waurþi*	*yrði*	*wurti*	*wurde*
	pl.	1	*waurþeima*	*yrðim*	*wurtīm*	*wurden*
		3	*waurþeina*	*yrði*	*wurtīn*	*wurden*
ptcp.			*waurþans*	*orðinn*	*(gi)wortan*	*worden*

With the exception of the 2 sg. in West Germanic, which probably continues the old perfect optative used as politeness form (**wurdīz*), the indicative singular goes back to PIE 1 sg. **u̯eu̯órta* (< **-h*$_2$*a* < **-h*$_2$*e*; cf. Skt. *vavárta* «I turned»), 2 sg. **u̯eu̯órt*^{s}*ta* (< **-th*$_2$*a* < **-th*$_2$*e*; Skt. *vavárttha*), 3 sg. **u̯eu̯órte* (Skt. *vavárta*). The other forms are based on the "weak" perfect stem **u̯eu̯r̥t-*, which was followed by accented endings in the parent language (cf., e.g., Skt. 1 pl. *vavr̥tmá*). Both the change in vocalism and the accent shift were faithfully preserved in Germanic, the former being seen in the substitution of PGmc. **-ur-* (< **-r̥-*) for **-ar-* (< **-or-*), and the latter in the "grammatical change" (*grammatischer*

Wechsel) of *-þ- to *-đ- – a feature lost in Gothic. The *-u- of the plural endings was probably generalized from the 3 pl. in *-un < *-n̥t, itself the replacement of an earlier r-ending (cf. Skt. vāvr̥túr < *-r̥s). The preterite optative rests on the PIE perfect optative, with *-ī- (< *-ih$_1$-), the variant of the mood sign originally proper to the plural and dual, generalized to the singular as well. The past passive participle of strong verbs points to a formation in *-anaz (var. *-inaz) akin to the Sanskrit perfect middle participle in -āná- (< *-onó-; cf. vāvr̥tāṇá-); before this suffix the root vowel *-u- was lowered to *-o- in North and West Germanic.

The inflection of strong verbs in Proto-Germanic was thus organized around four principle parts – the present, conventionally represented by the infinitive (*werþana^n); the preterite singular (1 sg. *warþ < *warþa); the preterite plural (*wurđum < *-ume (?)); and the past participle (*wođanaz). This system was maintained in all the medieval languages, but broke down in the early modern period, when the distinction between singular and plural stems in the preterite was generally abandoned. Traditional Germanic grammar recognizes six "classes" of ablauting strong verbs, of which the first five represent special cases of a single pattern. Class I consists of verbs with a root-internal i-diphthong: these originally had *-ei- in the present, *-oi- and *-i- in the preterite, and *-i- in the past participle, whence, e.g., Go. beitan : bait : bitum : bitans and OHG bīz(z)an : beiz : bizzum : (gi)bizzan. Class II, completely parallel, contains verbs with a u-diphthong (cf. Go. biudan : bauþ : budum : budans, OHG biotan : bōt : butum : (gi)botan). Class III, the type of *werþana^n, includes verbs with a root-internal sonorant that yielded -ur-, -ul-, -um- or -un- in the third and fourth principle parts; when the sonorant in question was a nasal, the present regularly had i-vocalism and there was no lowering of *-u- to *-o- in the past participle (cf. OHG bintan : bant : buntum : (gi)buntan «bind»). Classes IV and V consist of verbs with roots ending in a single consonant, either a sonorant (class IV; cf. Go. bairan : bar : berum : baurans, OHG beran : bar : bārum : (gi)boran) or an obstruent (class V; cf. Go. giban : gaf : gebum : gibans, OHG geban : gab : gābum : (gi)geban). The *-ē- of the preterite plural in class IV was borrowed from class V, replacing *-u- (*burum < *(bhe)bhr̥-mé), which was analogically extended to the participle on the model of classes II and III. In class V itself the source of the *-ē- in *gēbum is unclear. An analogical origin has been suggested, starting, e.g., from forms like *sētum «we sat» (< *se-zd- ?) and/or *ētum «we ate» (< *h$_1$e-h$_1$d- or *h$_1$ēd-). The participle *geƀanaz shows the substitution, probably ancient, of e-grade for the inconvenient zero-grade *gƀ-.

Class VI, containing verbs built to roots of the form CaC-, with a-vocalism, stands apart from the other types. The nucleus of this category was made up of verbs with old o-grade presents (e.g., Go. malan : mol : molum : malans, OHG malan : muol : muolum : (gi)malan «grind», cf.

Lith. *málti* « id ») as well as verbs with etymological **-a-* (e.g., Go. *hafjan* : *hof* : *hofum* : *hafans*, OHG *heffen* : *huob* : *huobum* : *(gi)haban*). The **-ō-* of the preterite thus continues both PIE **-ō-* and **-ā-*. The closest comparanda are Latin perfects like *scābī* (pres. *scabō*) and *fōdī* (pres. *fodiō* « I dig »), with corresponding forms in Greek and Celtic.

Not all strong verbs were capable of ablaut. Presents containing root-internal **-ai-, *-au-, *-ar-, *-al-, *-am-* and **-an-* (i.e., the structural counterparts, with *a*-vocalism, of classes I-III) might have been expected to change **-a-* to **-ō-* in the preterite, thus producing forms of the type **hōit(a)* to **haitana*ⁿ « order », **stōut(a)* to **stautana*ⁿ « push » and **hōlđ(a)* to **halđana*ⁿ « hold ». Proto-Germanic, however, did not permit long vowels in this position, having shortened tautosyllabic **ō* before glides and sonorants to **a* by Osthoff's Law. The distinction between the present and preterite stems was therefore marked in an exceptional way, namely, by the retention of reduplication. Reduplication was also preserved in strong verbs with **-ō-* in the present (e.g., **wōpjana*ⁿ : **wewōp(a)* « cry ») and, by analogy to these, in verbs with **-ē-* in the present and **-ē-* or **-ō-* in the preterite (e.g., **lētana*ⁿ : **lelōt(a)* « let »). The PGmc. reduplicated preterites maintained their existence in Gothic, where they were inflected regularly (cf, e.g., *haihait, lailot*, pl. *haihaitum, lelotum*, opt. 3 sg. *haihaiti, lailoti*, ptcp. *haitans, letans*). Traces of similar forms survive also in Old Norse (e.g., *sera* « I sowed » < **sezō-* (cf. Go. *saiso*) and West Germanic (e.g., OE *heht* « called » (cf. Go. *haihait*). For the most part, however, North and West Germanic converted the reduplicating verbs to ordinary strong verbs of a new type ("class VII"), characterized by the replacement of **-a-* by **-e-* in the second and third principal parts. Thus, the preterite of OHG *stōzan* (< **stautan*) was remade to *stioz* (< **steut*), while that of OHG *heizan* (< **haitan*) became *hiaz* (< **hē₂t* < **heit*;[14] cf. OE *hēt*, ON *hét*). The vowel **-ē₂-* was also consistently introduced into the preterite of verbs with *ē*-vocalism (cf. OHG *liaz*, OE *lēt*, ON *lét* (: PGmc. **lētana*ⁿ)); the same process took place dialectally in verbs of the structural type **halđan* (cf. OHG *hialt* vs. ON *helt* (**-ĕ-*, OE *hēold* (**-eu-*)). The genesis of these new preterites is hotly debated. It seems likely, however, that their nucleus was supplied by abridged reduplicated forms (cf. OE *wēpan* « cry », pret. *wēop* < **weup* < **wew(ō)p*, like *heht* < **heh(ai)t*) or, alternatively, by reduplicated forms with contraction (cf. ON *auka* « increase », pret. *jók* < **euk* < **e-auk* (= Go. *-aiauk*)). The hypothesis of a source distinct from the reduplicated perfect, much favored in the older literature, is hardly likely.

14. The "neo-diphthong" **ei* is not, of course, to be confused with PIE **ei*, which yielded **ī* in Germanic.

3.2.2. Weak verbs

Like every IE language, Proto-Germanic had a variety of derived present types, including causatives, iteratives, duratives and other semantically marked formations. These went back to well-known prototypes in the parent language, from which they inherited the property of being exclusively presential, i.e., of lacking associated aorist and perfect stems. Germanic was thus unable to furnish ablauting or reduplicated preterites to its secondary presents. Instead, it equipped them with preterites of a new type, the chief characteristic of which was a dental consonant, usually **-đ-*. Such preterites, and the verbs that formed them, are traditionally called "weak". Three main classes are recognized[15] :

Class I : the present stem originally contained the thematic suffix **-éi̯e/o-*, which gave first **-iji-/*-ija-* and then, by a kind of converse of Sievers' Law (cf. p. 266), **-ji-/*-ja-* after light syllables. In a small minority of cases, which may once have been strong (e.g., **sōk(i)jana*^{n} «seek», Go. *sokjan*), the present suffix was **-i̯e/o-*, which merged with **-éi̯e/o-* by becoming **iji-/*-ija-* after heavy syllables. The difference between the two treatments is clearest in Gothic : cf. *nasjan* «save», 3 sg. *nasjiþ* vs. *hausjan* «hear», 3 sg. *hauseiþ*, like *harjis* vs. *hairdeis*. North and West Germanic lost **-j-* before **-i-* ; instances of **-j-* that remained caused gemination in West Germanic (cf. OS *tellian* «say» vs. 3 sg. *teliđ*, OHG *zellen, zelit*). The present inflection was otherwise virtually identical with that of strong verbs. The nucleus of class I consists of PIE causatives, iteratives and (usually factitive) denominatives (cf., e.g., Go. *satjan*, OS *settian*, OE *settan* «set, cause to sit» (= Skt. *sādáyati*), Go. *fulljan*, OE *fyllan* «fill» (: PGmc. **fullaz* «full»).

Class II : the original suffix was PIE **-āi̯e/o-* (< **-eh*$_2$*-i̯e/o-*), which contracted to **-ō-* in Proto-Germanic. From a descriptive point of view, PGmc. **-ō-* thus replaced the thematic vowel in all present forms (cf. Go. *salbo, -os, -oþ*, etc., opt. 3 sg. *salbo*, inf. *salbon* «anoint»). In Old English the *a*-variant of the thematic vowel was restored through the interposition of a secondary **-j-*, borrowed from class I (cf. OE 3 sg. *sealfaþ* < **-ōþ* vs. 1 sg. *sealfie* < **-ōju*, inf. *sealfian* < **-ōjan*). Old High German and Old Saxon remade the 1 sg. in PGmc. **-ō* to *-ōm* (> OS *-on*) on the model of the irregular West Germanic verb **dōn* «do, put», where the 1 sg. **dōm* (OHG *tuom*, OE *dōm*), represented pre-Gmc. **dhedhōmi* (earlier **dhedhēmi*). Class II verbs, like their Latin counterparts of the first conjugation, continued a variety of inherited formations, with

15. The Proto-Germanic status of the Gothic fourth class (type *ga-fullnan* «be fulfilled», pret. *ga-fullnoda*) is uncertain.

denominatives greatly preponderant (cf. OHG OE *sealf*, OHG *salba* (*ō*-stem) « ointment »).

Class III : the recurring formal characteristic is a 3 sg. in **-aiþ*, with an associated 2 sg. in **-ais* and 2 pl. in **-aiþ* (cf. Go. *habaiþ* « has », OHG *habēt*, OE *hæfeþ* ; ON *vakir* « is awake »). Otherwise the individual languages show little agreement. Old High German has a regular paradigm, doubtless secondary, in *-ē-* (3 pl. *habēnt*, inf. *habēn*, etc.) ; Gothic and (usually) Old Norse intersperse forms in **-ai-* with forms in **-a-* (Go. 3 pl. *haband*, inf. *haban* ; ON 3 pl., inf. *vaka*) ; Old English and Old Saxon, where the type is vestigial, have **-ja-* where Gothic has **-a-* (OE 3 pl. *habbaþ*, inf. *habban*, OS *habbiađ, habbian*). Most class III verbs are stative or inchoative, and thus invite comparison with the Latin statives in *-ēre* (e.g., *tacēre* « be silent ») and Balto-Slavic statives in **-ē-* (e.g., Lith. *budė́ti* « be awake »). Traditional opinion favors a derivation from PIE **-ē-* (< **-eh₁-*), **-ēi̯e/o-* or **-əi̯e/o-* (< **-h₁-i̯e/o-*), none of which, however, can easily account for the attested forms. An alternative view takes the 3 sg in **-aiþ* from a 3 sg. middle in **-ai* < PIE **-oi* (cf. Skt. *duhé* "milks"), with **-þ* added from the other present classes.

The past participle of weak verbs was made with the PIE suffix **-tó-*. In the case of a few primary verbs this morpheme was added directly to the root (cf., e.g., ON *sóttr* « sought », OHG *(gi)suoht*, OE *sōht* < PGmc. **sōhtaz*) ; normally, however, it was substituted for **-i̯e/o-* in the pre-Germanic present stem (cf. Go. *nasiþs*, OHG *(gi)nerit*, OE *nered* < PGmc. **naziđaz* < **nose-tó-s*, like Lat. *monitus* « warned » < **mone-to-* ; Go. *salboþs*, OHG *(gi)salbōt*, OE *sealfad* < **salƀōđaz* < **solpā-tó-s*, like Lat. *cantātus* « sung »). Class III verbs interpose **-ai-* before the dental in Gothic and Old High German (*habaiþs, (gi)habēt* « had »), but OS *(gi)habd* and OE *hæfd* (< **haƀđaz*) probably preserve an earlier state of affairs.

The origin of the Germanic weak preterite is perhaps the most thoroughly studied morphological problem in any branch of the IE family. There is no generally accepted explanation of this category. In its Proto-Germanic form, the weak preterite was made by adding endings, wholly or partly distinctive, to a stem ending in a dental and closely resembling the stem of the past participle (**sōht-*, **nazid-*, **salƀōđ-*, **haƀđ-*). The reconstruction of the endings is problematic. In the plural and optative most of the languages have the same terminations as in the strong preterite (cf. e.g., ON *sóttum*, *-uð*, *-u*, opt. 3 sg. *sótti* ; OHG *suohtum*, *-ut*, *-un*, opt. 3 sg. *suohti*). Gothic, however, inserts an element *-ed-* before the normal endings, producing forms in *-(d)edum*, *-(d)eduþ*, *-(d)edun*, *-(d)edi*, etc., the antiquity of which is disputed. Of the singular paradigm, only the first person in **-(đ)ō*^{n} can be reconstructed with certainty (cf. Go. *-da*, Prim. Norse *-do* (*tawido* « I made » ; later ON *-ða*), OHG *-ta*, *-de*).

The 3 sg. is often assumed to have anded in PGmc. **-(đ)ē*, largely on the strength of Prim. Norse *-de* (*tawide* «(he) made», later ON *-ðī*), but this is phonologically doubtful and contradicted by the still earlier runic form *talgidai* «(he) carved», apparently < PGmc. **-đai*. The 2 sg. forms are flatly incompatible : Gothic has *-(d)es*, Old High German and one dialect of Old Saxon point to **-(đ)ōs* (OHG *-tōs*), while the other languages favor **-(đ)ais* (ON *-ðir*, OE *-des*, OS (one dialect) *-đes*, *-đas*). Many attempts have been made to interpret these facts. Adherents of the "composition theory" have explained the weak preterite as a periphrastic formation containing the verb "to do" (WGmc. **dōn*). The strongest argument in favor of this analysis comes from the Gothic forms in *-dedum*, etc., which recall OHG *tātum*, *-ut*, *-un* «we (you, they) did»; the strongest argument against it comes from the vocalism of the singular endings, and of the 3 sg. in particular. Other, less attractive periphrastic explanations have been proposed, some involving the past participle itself. Probably the most plausible alternative to the *do*-theory remains the view, due to Collitz, that takes the weak preterite from a nucleus of inherited perfect middles in 3 sg. **-(đ)ai* < pre-Gmc. **-(t)oi*.

Among the scattered archaisms that survived the restructuring of the PIE verbal system in Germanic, special mention must be made of the "preterito-presents", a small but important Proto-Germanic class that partook of the properties of both strong and weak verbs. These were in origin PIE or pre-Germanic perfects which resisted the normal semantic evolution of the perfect and retained their value as stative presents. From a Germanic point of view, the synchronic presents of the preterito-presents were thus formally similar to strong preterites, while their synchronic preterites and past participles were weak. More than a dozen such verbs occur in Gothic, including *witan* «know» (1 sg. *wait* : pl. *witum*, pret. *wissa*; cf. Skt. *véda* : *vidmá*; Gk. *oîda* : *ídmen*), *ga-daursan* «dare» (*ga-dars* : *ga-daursum*, pret. *ga-daursta*; cf. Skt. *dadhárṣa* : *dadhr̥̥ṣmá*) and *munan* «think» (*man* : *munum*, pret. *munda*; cf. Gk. *mémona* : *mémamen*). Comparable forms occur in the other early dialects, but like ablauting verbs generally, the preterito-presents tended to decline in the later history of Germanic. In the modern languages they survive mainly as auxiliaries (cf. Eng. *shall, can, may*; NHG *sollen, können, mögen*), joined in this function by another "irregular" verb, the old athematic present optative **wilī-* «wish, want» (Go. 3 sg. *wili*, 1 pl. *wileima*; cf. Eng. *will*, NHG *will, wollen*).

3.3. Syntax

The syntax of Proto-Germanic was typologically similar to that of the other early IE languages, especially Italic, Greek and Balto-Slavic. It will suffice here to note a few special features.

Agreement of adjectives with their head nouns, and pronouns with their antecedents, in number, gender and case was a general IE feature, but it was a peculiarity of Proto-Germanic that plural noun phrases of mixed gender induced *neuter* agreement (cf., e.g., Go. *jah qemun þan aiþei is jah broþrjus is jah uta standandona* [nt. pl.] *insandidedun du imma, haitandona* [nt. pl.] *ina* «and his mother and brothers came and, standing outside, sent in to him, calling him» (Mk. 3.31)). This rule had its origin in the accidental merger and confusion of the old masculine dual ending *-ō < PIE *-o-h_1, which became obsolete in Germanic, with the productive neuter plural ending *-ō < *-ā < *-e-h_2.

Germanic was relatively conservative in its treatment of the PIE prepositional adverbs (e.g. *pro «in front (of)», *eni «inside (of)», *ko(m) «together (with)», etc.). As in the classical languages, these had a twofold development, yielding prepositions on the one hand and proclitic verbal prefixes on the other. The univerbation of the new prefix + verb complexes, however, was not yet complete in Proto-Germanic, as can be seen from occasional instances of tmesis like Go. *uz-uh-hof* «and (*-uh-*) lifted up» (*us-hafjan* «lift up») and *ga-u-ƕa-seƕi* "whether (*-u-*) he might see anything (*-ƕa-*)" (*ga-saiƕan* «see»). Such cases of "Wackernagel's Law" – the movement of enclitics to second position in the sentence – are wholly confined to Gothic. Old Norse abandoned the older type of verbal composition entirely, while the West Germanic languages carried the process of fusion to completion (cf. NHG *erheben* = Go. *us-hafjan*). All the Germanic languages show a later tendency to create new compound-like phrases ("separable compounds") using stressed adverbs of place or direction – the type of Eng. *go in*, ON *ganga inn*, NHG *(hin)eingehen*, etc.

Proto-Germanic word order was "free" in the same sense as in other highly inflected IE languages; that is, a change in the order of the constituents in a given sentence often affected its emphasis or focus, but not its grammaticality. There is reason to believe that in the earliest Germanic dialects the basic, "unmarked" word order in both main and subordinate clauses was *subject – object – verb* (SOV), as is still the case in the earliest Primitive Norse inscriptions. Even in Proto-Germanic, however, main clauses freely employed the alternative order *subject – verb – object* (SVO) to focus attention on the direct object. SVO order predominates in main clauses in the early literary languages and their medieval and modern descendants, in some of which (e.g., English) it eventually spread to subordinate clauses (cf. *The man saw the woman; when the man saw the woman...*), while in others (notably German) it contrasted with SOV order in subordinate clauses (cf. *Der Mann sah die Frau; als der Mann die Frau sah...*). Relative clauses, marked only by a particle in Proto-Germanic (cf. Go. *sa* «he, that one», *saei* «(he) who»; ON *sá* : *sá er*; OHG *der* : *der dar*; OE *se* : *se þe*), patterned in this respect like other subordinate structures. Subject-verb inversion, generating sentences with VS(O) order,

was obligatory in main clauses introduced by sentence-connective particles like ON *þá*, OE *þā* and OHG *da* (cf. ON *þá gekk hann inn* «then went he in», etc.). Such "verb second" or "V2" effects have been explained as an indirect reflex of Wackernagel's Law, since finite verbs in main clauses were enclitic in Proto-Indo-European.

The synchronic study of Germanic syntax – relativization, topicalization, inversion, and "movement" generally – is a subject of great current interest in theoretical linguistics. If it proves possible to advance from synchronic to diachronic insights, the rich evidence of the Germanic languages may yet play as important a role in the development of a mature theory of historical syntax as it once played, in the time of the Neogrammarians, in the development of a theory of historical phonology.

Bibliographie

BACH A. (1965). – *Geschichte der deutschen Sprache*[5], Heidelberg, Quelle und Meyer.

BAMMESBERGER A. (1986). *Der Aufbau des germanischen Verbalsystems*, Heidelberg, C. Winter.

BAUGH A. (1978). *A history of the English language*[3], Englewood Cliffs, Prentice-Hall.

FOURQUET J. (1938). – *L'Ordre des éléments de la phrase en germanique ancien*, Paris, Belles Lettres.

– (1956). – *Les Mutations consonantiques du germanique; essai de position des problèmes*, Paris, Belles Lettres.

HAUGEN E. (1976). – *The Scandinavian languages : an introduction to their history*, Cambridge, Mass, Harvard University Press.

JOLIVET A. and MOSSÉ F. (1947). – *Manuel de l'allemand du Moyen Age, des origines au* XIV[e] *siècle*, Grammaire. Textes. Glossaire. Nouvelle édition. Paris, Aubier.

KRAHE H. (1966-1967). – *Germanische Sprachwissenschaft*[6]. I. Einleitung und Lautlehre. II. Formenlehre. III. Wortbildungslehre, von W. Meid, Berlin, De Gruyter.

KRAUSE W. (1971). – *Die Sprache der urnordischen Runeninschriften*, Heidelberg, C. Winter.

LEHMANN W. (1986). – *A Gothic etymological dictionary*, with bibliography prepared under the direction of Helen-Jo J. Hewitt, Leiden, E.J. Brill. Based on Sigmund Feist, *Vergleichendes Wörterbuch der gotischen Sprache*[3] (1939).

LOCKWOOD W.B. (1965). – *An informal history of the German language*; with chapters on Dutch and Afrikaans, Frisian and Yiddish, Cambridge, England, W. Heffer.

MEILLET A. (1930). – *Caractères généraux des langues germaniques*[4], Paris, Hachette.

MOSSÉ F. (1958). – *Esquisse d'une histoire de la langue anglaise*[2], Lyon : I.A.C.

– (1942). – *Manuel de la langue gothique*, Grammaire. Textes. Glossaire, Paris, Aubier.

NIELSEN H. (1989). – *The Germanic Languages – Origins and Early Dialectal Interrelationships*, Tuscaloosa, London, University of Alabama Press.

RAMAT P. (1981). – *Einführung in das Germanische*, Tübingen, Niemeyer.

SEEBOLD E. (1970). – *Vergleichendes und etymologisches Wörterbuch der germanischen starken Verben*, The Hague, paris, Mouton.

WEERMAN C. (1989). – *The V2 conspiracy : a synchronic and a diachronic analysis of verbal positions in Germanic languages*, Dordrecht, Providence, Foris.

WESSEN E. (1968). – *Die nordischen Sprachen*, Berlin, De Gruyter.

CHAPITRE 14

L'ITALIQUE

Enrico CAMPANILE

Si nous prenons comme point de repère chronologique les derniers siècles de l'Antiquité, nous voyons que l'Italie était caractérisée par une fragmentation linguistique très marquée. Dans certains cas, il s'agit de langues importées à une époque relativement récente, et limitées à des établissements coloniaux très restreints sur le plan géographique (grec et phénicien) ; celles-ci resteront en dehors de notre exposé, puisqu'elles ne diffèrent pas de manière significative des parlers respectifs de leur patrie originelle. Mais dans la plupart des cas, il s'agit de langues que nous pouvons qualifier d'indigènes, parce que les sujets parlants concernés s'étaient établis en Italie depuis fort longtemps, au moins dès le IIe millénaire av. J.-C. Les parlers gaulois occupent une position intermédiaire, car leur arrivée en Italie est plutôt récente (certainement pas antérieure à la colonisation grecque et phénicienne), mais les modalités de leur établissement, sur un territoire vaste et continu et avec l'élimination systématique de presque toutes les langues préexistantes, sont identiques à celles des langues indigènes.

Certaines de ces langues n'étaient pas indo-européennes, c'est le cas de l'étrusque ; d'autre langues, comme le rhétique, ont une classification controversée. Cependant, dans la plupart des cas, elles appartiennent à la famille indo-européenne.

Leur définition dialectologique n'est toutefois pas facile, puisque nous ne les connaissons que d'une manière plutôt partielle (à l'exception du latin, qui reste en dehors de notre exposé), grâce à un nombre limité de textes épigraphiques (dont l'interprétation n'est pas toujours certaine), ainsi qu'à un nombre plus limité encore de gloses conservées chez des auteurs grecs ou latins.

Pour beaucoup de langues donc, l'analyse dialectologique est objectivement impossible, que ce soit à cause de l'insuffisance de matériaux ou de la carence d'interprétations linguistiques et textuelles convaincantes qui en découle ; quoiqu'avec des réserves et des lacunes, on ne peut tenter d'en établir une que pour les dialectes italiques (osque, ombrien et dialectes mineurs à savoir les parlers des Volsques, des Marses, des Vestiniens, des Marrucins, des Péligniens) et le vénète, qui offrent des *corpora* épigraphiques d'une certaine consistance et qui font l'objet d'études sérieuses depuis de nombreuses années.

1

Pendant longtemps, la définition dialectologique des langues de l'Italie antique consistait essentiellement à reconstruire les unités généalogiques intermédiaires dont la scission aurait permis à chaque langue de se constituer. De ce point de vue, le modèle de l'italo-celtique proposé par Lottner en 1861 demeure exemplaire : celui-ci aurait engendré d'une part le celtique commun et de l'autre l'italique commun, ancêtre du latin et des dialectes italiques. Cette hypothèse généalogique a été détruite en 1929 par Marstrander et par Devoto, qui, indépendamment l'un de l'autre, ont démontré l'absence d'innovations anciennes et communes à l'italique et au celtique qui seraient nécessaires pour garantir la notion d'italo-celtique. Mais Devoto est allé bien plus loin, dans la mesure où il niait également l'hypothèse d'un italique commun, estimant que les éléments communs au latin et aux dialectes italiques étaient la conséquence d'une convergence tardive, tandis que les divergences seraient, elles, anciennes et profondes.

Mais, alors que le refus de l'unité italo-celtique a trouvé une approbation substantielle (peut-être aussi parce qu'il provenait d'un grand celtiste tel que l'a été Marstrander), le refus de l'unité italique a connu bien moins de succès ; de nombreux chercheurs continuent même d'admettre que cette unité généalogique intermédiaire existe et vont jusqu'à y rattacher d'autres langues de l'Italie antique, comme le vénète (Beeler, 1949) ou le sicule (Vetter, 1962).

Rome a ses premiers contacts avec le monde italique au Ve siècle av. J.-C., et, à partir du siècle suivant, elle commence progressivement à imposer son hégémonie ; sur le plan linguistique, ce fait aboutit à une romanisation totale des dialectes italiques (I^e^ siècle av. J.-C.). Mais un tel processus historique ne peut certainement pas s'identifier avec la convergence dont parle Devoto ; en réalité, nous ne voyons ici aucune convergence mutuelle, mais seulement un échange linguistique qui laisse en substance inaltérée la langue victorieuse.

Pendant une période très longue et indéfinie, mais antérieure au Ve siècle av. J.-C., la situation est toutefois bien différente : Rome est une modeste et petite réalité, sans rapports avec le monde italique. Dans cette phase, qui remonte très loin avant la naissance officielle de Rome, il est naturel que l'on observe une phase progressive de divergence entre le latin et les dialectes italiques, puisqu'il n'y avait pas de rapports historiques entre eux.

Mais, d'autre part, il existe entre ces deux groupes une indéniable affinité linguistique que l'on ne peut situer que dans une phase pré-ethnique, pour les raisons qui ont été exposées précédemment. L'affirmation de Beeler (1966), selon laquelle les similitudes entre le latin et les dialectes italiques peuvent s'expliquer aussi bien par une « convergence au sein de l'aire linguistique de l'Italie péninsulaire à une époque préhistorique », que par l'hypothèse d'une unité italique est inacceptable. Au contraire, nous ne voyons aucune possibilité de convergence en Italie d'un point de vue préhistorique, à moins que l'on ne suppose de très anciens contacts inter-ethniques qui n'auraient de fondement que dans les affinités linguistiques ; il s'agirait alors d'un véritable cercle vicieux.

Le moment de contacts intenses qui justifient l'étroite contiguïté entre les dialectes italiques et le latin, doit donc se situer dans une zone géographique extérieure à l'Italie, et concerner une période historique assez proche de la dissolution du monde indo-européen : quant à savoir si nous voulons voir ensuite dans ce moment une unité intermédiaire au niveau généalogique ou une phase de contacts très intenses qui font que deux langues différentes à l'origine deviennent semblables entre elles, ceci est un problème substantiellement nominaliste que tout chercheur essaie de résoudre dans le cadre de sa vision générale de l'histoire de l'indo-européen. Le fait essentiel, c'est de reconnaître les trois phases fondamentales après la phase communautaire indo-européenne : la première, pré-ethnique, de voisinage ; la deuxième, en partie pré-ethnique mais aussi ethnique, d'éloignement ; la troisième, exclusivement ethnique et située dans la période historique, de dissolution de l'italique dans le latin. Cette segmentation nous permet ainsi de sortir du pur dilemme généalogique, et l'hypothèse d'une unité intermédiaire est ramenée à un niveau plus raisonnable, c'est-à-dire comme un événement possible parmi tous ceux qui caractérisent l'histoire d'une langue.

2

Une unité ancienne (même si on l'entend dans un sens non généalogique) doit laisser des traces dans des faits de grande importance structurale et non pas dans des éléments de surface qui peuvent être facilement remplacés ou renouvelés.

Si nous considérons maintenant le système verbal italique et le système verbal latin dans leurs structures fondamentales, on remarque tout de suite que leur identité absolue est d'ordre innovateur.

Tous les thèmes du présent (mis à part quelque fossiles, comme «être») peuvent être divisés en quatre classes :

verbes en *ā* : osque [ci-après : o.] *deivatud* (comme lat. *iurato*),
verbes en *ē* : o. *licitud* (comme lat. *liceto*),
verbes en *e/o* : o. *actud* (comme lat. *agito*),
verbes en *ī* : ombr. *seritu* (comme lat. *servito*).

Dans d'autres langues aussi on trouve de nouvelles structurations du verbe par rapport à la situation indo-européenne; mais l'exemple que nous venons de donner est propre au latin et aux dialectes italiques; en germanique et en irlandais, par exemple, la classification est complètement différente.

L'ensemble du système verbal latin et italique se base, en outre, sur l'opposition entre *infectum* et *perfectum* (cf. Porzig, 1960). Nous transcrivons ici le schéma latin de Leumann, 1977, 507 :

Zeitstufe	Praesenssystem akt. (auch pass.) Praesensstamm *flē-*			Perfektsystem (akt.) Perfektstamm = *flēv-*		
	Indik.	Konj.	Infin.	Indik.	Konj.	Infin.
Gegenwart	*flē-s*	*fle-ā-s*	*flē-re*	*flēv-istī*	*-er-is*	*isse*
Vergangenheit (Praeteritum)	*flē-bā-s*	*flē-rē-s*	—	*-erā-s*	*-issē-s*	—
Zukunft (Futurum)	*flē-bi-s*	—	périphr.	*-er-is*	—	périphr.

On peut parfaitement l'utiliser pour les dialectes italiques en substituant les formes indigènes aux formes latines; la pauvreté de la documentation nous oblige, naturellement, à ne pas utiliser les formes d'un même

verbe qui devraient être reconstruites pour la plupart :

INFECTUM	FAAMAT « vocat »	*portaia* « portet »	FATÍUM « dicere »
	FUFANS « erant »	FUSÍD « esset »	
	deiuast « iurabit »	______	
PERFECTUM	DEDED « dedit »	*fefacid* « fecerit »	*lexe* « legisse »
	?	?	
	BENUS « veneris »	______	

Comme on peut le remarquer, il manque dans les textes italiques les formes du passé du *perfectum*, que nous avons indiquées par « ? » ; mais cette absence ne dépend vraisemblablement que de la pauvreté de la documentation. Il faut, du reste, tenir compte du fait que d'autres catégories grammaticales de l'italique sont également attestées par un nombre très limité de formes.

Comme c'était déjà le cas pour la subdivision des thèmes du présent en quatre classes, on peut également observer, à propos de la restructuration du système verbal par rapport aux modules originels indo-européens, que le système représenté ci-dessus se retrouve seulement en latin et dans les dialectes italiques, bien que des tendances analogues puissent se retrouver ailleurs.

La création d'un système de formes nominales liées étymologiquement et fonctionnellement à chaque verbe est une autre caractéristique structurale qui leur est commune.

Le participe présent en **-nt-* (ombr. ZEŘEF : lat. *sedens*) et le participe passé en *-to-* (o. *scriftas*/lat. *scriptus*) sont des formations héritées, mais l'infinitif (o. MOLTAUM : lat. *multare*), le supin (ombr. ANZERIATU : lat. *observatum*) et le gérondif (o. SAKRANNAS : lat. *sacrandus*) ont un caractère nettement novateur. Quelques précisions s'imposent cependant.

La diversité morphématique qui distingue l'infinitif latin de l'infinitif italique ne peut être utilisée comme argument. Comme nous l'enseignent l'iranien et le védique, qui ont eux aussi développé un infinitif (Benveniste, 1935), on utilisait différents types de formations nominales pour l'infinitif (noms radicaux, en **-ti-*, en **-yo-* etc.), si bien que les deux morphèmes différents du latin et de l'italique représentent simplement le résultat de processus sélectifs qui ont eu lieu dans une phase post-unitaire. La généralisation du participe en **-to-* aux dépens de celui en **-no-/*-meno-*, qui n'a survécu que comme élément nominal (lat. *plēnus* et *alumnus* ne sont évidemment pas les participes de *plēre* et *alere*), ne devra pas s'expliquer différemment sur le plan morphologique.

Le supin et le gérondif sont parfaitement identiques aussi sur le plan morphologique en latin et dans les dialectes italiques. Le premier est un nom en **-tu-*, figé à l'accusatif dans sa fonction de mouvement vers le lieu

(ombr. ANZERIATU ETU = lat. *observatum ito*) ; le second a une étymologie controversée.

À l'origine, la nature nominale de l'infinitif n'avait rien à voir avec l'opposition actif-passif ; mais son insertion dans le système verbal entraîna l'attribution de la marque de l'actif, et par conséquent la création de son correspondant passif. Un infinitif présent passif n'est pas attesté dans les dialectes italiques ; mais il a dû certainement exister, puisque sa forme correspondante au passé est attestée : ombr. KURATU ERU = lat. *curatum esse*.

Un système asymétrique avec sept cas au singulier et quatre cas au pluriel se constitue pour le nom.

Singulier :
nom. m. HÚRZ, n. SAKARAKLÚM («jardin», «sacellum»)
gén. SAKARAKLEÍS
dat. SAKARAKLÚÍ
acc. m. HÚRTÚM
voc. m. SILIE
abl. SAKARAKLÚD
loc. SAKARAKLEÍ

Pluriel :
nom. m. NÚVLANÚS, n. *comono* («Nolani», «comices»), voc. ombr. IKUVINUS
gén. NÚVLANÚM
acc. m. NÚVLANÚSS (ombr. VITLUF «veaux»)
dat. abl. loc. NÚVLANÚÍS

À ce schéma (osque, avec très peu de formes reconstruites) correspond parfaitement le paradigme latin, que nous reportons sous sa forme archaïque : sing.nom. *lupos*, gén. *lupī*, dat. *lupōi*, acc. *lupom*, voc. *lupe*, abl. *lupōd*, loc. *lupoi* ; plur. nom. voc. *lupoi*, gén. *lupom*, acc. *lupōs*, dat. abl. loc. *lupois*.

De ce point de vue structural, la divergence morphologique entre SAKARAKLEÍS et *lupī* n'a pas d'importance : il s'agit, dans les deux cas, d'innovations par rapport à l'**-osyo* originel, dont le latin conserve une seule attestation dans *popliosio valesiosio* du *Lapis Satricanus*. Ce qui compte ici, en réalité, c'est l'identité des deux systèmes.

À côté de ces convergences de caractère très général, on a une série d'éléments communs beaucoup plus ponctuels. Nous en relèverons quelques-uns ici.

À côté du subjonctif en *-ā-*, dont il sera question dans le chapitre sur le celtique, le latin et les dialectes italiques possédaient un subjonctif en *-ē-*, limité à la première classe de verbes : o. *deivaid* «qu'il jure» (< **deiwāyēt*), lat. *iuret* (< **yousāyēt*).

Les deux groupes possèdent en outre un subjonctif imparfait en -*sē*- : lat. *foret*, o. FUSÍD, tous deux issus de **bhusēt*. Que le gall. *cerhynt* « amassent » puisse remonter à **karasēnti* est une hypothèse qui ne trouve pas de fondement dans le cadre du paradigme gallois ou britannique.

Des formes anciennes d'optatif athématique ont été transférées dans la catégorie du subjonctif (lat. *sient*, ombr. *sins*, o. *-sins*) ; des traces ultérieures survivent encore dans le lat. *faxint*, *capsint* et d'autres semblables. En rapport avec ces innovations, le subjonctif originel **esed* est passé au futur (*erit*), tandis que les dialectes italiques utilisaient la forme supplétive **bhu-se-t* (o. ombr. *fust*), avec un type morphologique qui n'est pas étranger au latin (*faxō*, *capsō*, etc.).

3

Sur cette base d'innovations anciennes et communes, une série d'innovations divergentes ont eu lieu dans les deux groupes, et leur justification historique est à rechercher, comme nous l'avons vu, dans une longue phase de séparation ethnique. Nous en dressons une petite liste.

Des formations d'origines diverses et assez récentes pour séparer l'osque de l'ombrien confluent dans le parfait latin et italique ; mais d'un point de vue morphologique, deux tendances nettes et divergentes se manifestent. Le latin, en effet, tend à généraliser les désinences du parfait indoeuropéen : *dedī*, *dedistī* et *dedere* sont fondamentalement comparables au véd. *dade*, *dadātha* et *dadur*, là où les dialectes italiques généralisent les désinences secondaires de l'aoriste : o. MANAFUM « mandavi », PRÚFATTED « approuva », UUPSENS « firent » (**-m*, **-t*, **-nt*). Mais, d'autre part, l'italique conserve l'ancien participe du parfait (o. *sipus* qui traduit le lat. *sciēns*, et *facus* « qui a fait »), tandis que le latin n'en a qu'un petit résidu dans l'adjectif *memor*.

Le latin *sunt* a éliminé l'ancienne désinence athématique (cf. got. *sind*, v. irl. *it*, gr. myc. *eesi*) en faveur de la désinence thématique et le même phénomène s'est produit en partie dans le participe, où l'ancien *-sēns* (dans *praesēns*) coexiste avec le nouveau *sōns*, tandis que dans les dialectes italiques seules des formes conservatrices sont attestées : o., ombr. *sent* « ils sont », o. *praensentid*.

Dans le domaine des pronoms, les formes italiques *tiom* « toi » et *siom* « soi », comparées au lat. *te(d)* et *se(d)*, permettent de reconstruire une chronologie relativement précise :

(a) i.-e. **twĕ*, **swĕ* (gr. σε, pamphyl. Ϝhε) ;
(b) lat. ital. **tē*, **sē* (avec l'allongement des monosyllabes à terminaison vocalique),

(c1) en italique, redétermination par le morphème nominal *-om*;
(c2) en latin, introduction du double *tēd* et *sēd* par analogie avec le couple abl. *tē/tēd*, *sē/sēd* (Pfister 1977, 202).

Le lat. *sibi* est analogue à *tibi* : l'ombr. *seso* «à soi», si l'on pense qu'il vient de *swoi-so – ce qui est loin d'être sûr –, conserverait alors sa forme ancienne (cf. cypr. Ϝοι).

L'italique conserve l'adjectif pronominal *k^{w}o-teros «lequel des deux» (véd. *katara-*), que le latin remplace par *uter*, mais il remplace l'ancien *k^{w}etrwr̥s «quatre fois» (véd. *catur*, av. *caθruš*, lat. *quater*) par *petiropert*, d'après un type composé qui, à l'origine, n'était valable que pour les numéraux supérieurs à quatre.

Dans la morphologie nominale, l'italique conserve le nominatif pluriel en *-ōs* (o. NÚVLANÚS), le latin le remplace par la désinence pronominale *-oi*; on lisait déjà dans le *Lapis Satricanus* le nominatif *mamertoi* (Campanile, 1985); la même chose se produit pour les thèmes en *-ā*.

Dans les sites historiques se réalise enfin le moment de la dissolution des dialectes italiques. Comme Vendryes l'avait déjà observé à propos d'un dialecte breton, cette dissolution se fait de l'intérieur : l'écroulement final est très rapide précisément parce que si l'apparence formelle restait italique, le latin fournissait désormais les véritables modèles culturels et linguistiques. Des syntagmes comme *comono hafiest* «tiendra la réunion», *zicolom dicum* «intenter un procès», *egmo toutico* «l'état» sont osques dans leurs éléments formels, mais ils demeureraient historiquement incompréhensibles sans la pré-existence du modèle lat. *comitia habere*, *diem dicere*, *res publica*. Dans cet exemple, il est manifeste que le groupe latin a remplacé le *touto* indigène.

Cette pression du latin peut également se manifester dans des détails minimes, mais ce n'est pas pour autant qu'ils sont insignifiants. Dans une inscription osque de Schiavi d'Abruzzo (Po. 34) nous lisons ***úpsed g. paapii.g.f.*** «(le) fit Gaius Papius fils de Gaius» : mais cette formule est la formule onomastique romaine, et non osque, qui disposait différemment les éléments et n'employait pas le lexème «fils».

La même inscription commence par les mots ***m.t.ni dekitiúd***, c'est-à-dire avec un ablatif absolu («étant *meddix* Numerius Decitius») qui représente un hapax syntaxique en osque et qui ne peut se justifier que comme latinisme syntaxique.

Cette hégémonie culturelle du latin ira jusqu'à porter atteinte à la forme même des parlers italiques, et cette phase est documentée par des textes comme la *Tabula devotionis* de Cumes qui est publiée aussi bien dans le *Corpus Inscriptionum Latinarum* (I^2 1614) que dans le recueil italique de Vetter (n°7). L'inévitable étape finale est l'acceptation du latin.

4.

Le latin et les dialectes italiques dont nous avons jusqu'à présent parlé ne constituent en réalité que les pôles extrêmes et les plus connus d'une réalité linguistique beaucoup plus complexe, puisqu'en Italie centrale nous trouvons d'autres témoignages linguistiques, que l'on a différemment insérés du point de vue généalogique dans le cadre d'une unité italique.

Sur les inscriptions du Sud du Picenum, éditées récemment (Marinetti, 1985) et qui doivent être encore longuement analysées, tout discours serait prématuré. Mais on peut désormais exclure que le falisque doive être considéré comme une variante locale du latin, avec des influences italiques ou moins tardives.

En réalité, le falisque présente une véritable autonomie non seulement par rapport au latin mais aussi aux dialectes italiques, et si cette autonomie apparaît quantitativement modeste, c'est à cause de la pauvreté de la documentation, mais elle n'est pas pour autant moins évidente.

Le cas du futur *carefo* « je me passerai (de) », *pipafo* « je boirai » est exemplaire : ce ne sont pas des formes italiques, parce que l'italique ne connaît pas de futur en **-bh-*, et ce ne sont pas des formes latines, parce que le latin possède certes l'aboutissant *-b-* de **-bh-* mais ne possède pas le présent à redoublement **bibare* « boire », la coïncidence phonétique avec l'osque et morphologique avec le latin est par conséquent non signifiante.

On ne doit donc pas penser au falisque comme à une espèce de *Mischsprache* qui tire son individualité de la présence simultanée de traits dont chacun – mais non pas tous – se retrouve dans un dialecte voisin. En effet, il y a également des éléments qui le séparent nettement aussi bien du latin que des dialectes italiques. C'est le cas de *foied* (**ĝho-diēd*) « aujourd'hui », dont le consonantisme initial s'oppose au lat. *hodiē* et à l'italique (où ce lexème, s'il existait, commencerait nécessairement par *h-*). C'est le cas du pronom *ues* « vous » (lat. *uos*, o. *uus*), qui ne peut être comparé (mais, à notre avis, tout à fait par hasard : il s'agit de remaniements analogiques sur le vocalisme du singulier **mē*) qu'au celt. **swēs* (gall. *chwi*, vieil irl. *sí*). C'est le cas de *lecet* « git » qui, en tant que verbe, n'a des correspondants qu'en dehors de l'Italie (gr. λέχεται, got. *ligan* etc.).

Cette individualité est encore plus évidente si l'on considère des paradigmes entiers. Au parfait, par exemple, la première personne du singulier continue le parfait moyen européen : *peparai* « procurai » comme le lat. *peperī* (et le véd. *dadé*), mais la troisième personne du pluriel continue la désinence de l'aoriste : *f(if)iqod* « formèrent » (**-nt*). Il ne serait pas fructueux de voir dans le premier cas une coïncidence significative avec le latin, et dans le second avec les dialectes italiques. En réalité, des mélanges de ce genre se produisent même ailleurs, sans aucune implication

généalogique : le vieil irl. *cechnitir* « chantèrent » arrive à réunir dans le morphème *-itir* la désinence indo-européenne de l'aoriste et la désinence du parfait (*-nt* + **r*).

D'autre part, le falisque ne manque pas d'éléments qu'il partage parfois avec le latin, parfois avec l'italique, mais rien ne nous autorise à les exploiter d'un point de vue généalogique. Nous pouvons rappeler ainsi en premier lieu les lexèmes *hileo* « fils » et *filea* « fille » du falisque qui coïncident avec le lat. *fīlius* et *fīlia*, mais non avec l'italique **puklo-* et FUTÍR ; en deuxième lieu, le parfait à redoublement *fifiked* « forma », est comparable à l'o. FIFIKUS (fut. ant.), mais le latin possède l'ancien aoriste sigmatique *fīnxī*. Enfin, il est difficile d'évaluer la portée du subj. *douiad* « donne », qui vient du thème **dowī-* que nous retrouvons dans la forme composée de l'ombrien *purdouitu* « offre », mais qui apparaît également ailleurs (dans sa forme simple osque *dida*, ombr. *dirsa* « qu'il donne » < **di-dāt*),

Il y a enfin des localités extra-urbaines (Préneste, Pesaro, Tivoli, Spolète, etc.) dont la langue, bien qu'elle soit latine en substance, présente différents éléments qui la distinguent du modèle romain ; dans ces cas, nous pourrons donc parler de latin dialectal que l'on opposera au latin urbain. Rappelons ici quelques-uns de ces éléments :

a) monophtongaison des diphtongues : *coraveron* (Préneste, = *cura-verunt*), *Marte* (Tivoli, = *Marti*), etc.,

b) perte du *-s* même après voyelle longue : *matrona pisaurense* (Pesaro), etc. ;

c) syncope vocalique plus intense qu'à Rome : *dedro* (Pesaro, = *dederunt*), *merto* (Préneste, = *merito*), etc. ;

d) distribution de *h* et *f*, en tant qu'aboutissants de consonnes aspirées, différente par rapport au latin : *fircus* (attribué aux « Sabins ») contre *hircus*, etc. ;

e) dat. sing. en *-a* pour les thèmes en *-a* : *Fortuna* (Préneste), etc.,

f) gén. sing. en **-os* pour les thèmes consonantiques : *nationu* (Préneste), etc. ;

g) nom. pl. en **-eis* pour les thèmes en *-o* : *libereis* (Sora), *duomvires* (Cora), *magistere(s)* (Préneste), etc.

Il s'agit d'innovations qui sont d'époque archaïque, quand Rome n'avait pas encore imposé son hégémonie à toute la région du Latium. Ce type linguistique est donc indigène dans certaines zones (comme Préneste) ; dans d'autres comme Pesaro, il est dû à une colonisation provenant de territoires extra-urbains.

Il est toutefois nécessaire de souligner que la phase précédant le latin dialectal, tel qu'il est attesté aujourd'hui, n'est pas différente de celle qui a précédé le latin urbain, et que la naissance de divergences ne peut se rapporter à une phase pré-ethnique. Cela est également le cas pour les sonores aspirées ; le fait qu'au Ve siècle av. J.-C. **-bh-* n'était pas encore

devenu *f* en osque (démontré par Rix, 1957), est un indice important qui démontre que le phénomène est récent en latin aussi.

5

L'excellent manuel vénète de Lejeune (1974) donne la liste de plus de cinquante éléments qu'il convient de considérer pour déterminer la position linguistique de cette langue.

Ils ne sont pas, naturellement, tous de la même valeur. Le fait, par exemple, que le vénète continue à distinguer *o* de *a* n'est pas un argument suffisant pour le distinguer du germanique, puisque nous savons que les emprunts celtiques participent au passage de *o* à *a* (gall. *Volcae* : v.h.a. *Walha*), alors que ce n'est pas le cas pour les emprunts du latin (lat. *coquere* : v.h.a. *kochon*) ; cela signifie, en termes chronologiques, qu'en germanique le passage de *o* à *a* se produit entre 500 av. J.-C. et le début de l'ère chrétienne. Le vénète, quand il était encore vivant, avait le même vocalisme que le germanique.

On ne peut pas non plus admettre que le vénète ait reconstruit l'adjectif « mort » (**mr̥two-* à la place de **mr̥to-*) sur le modèle de « vif » (*g^{w}*īwo-*) comme c'est aussi le cas en latin (*mortuus* à la place de **mortus*, d'après *vivus*) et soutenir en même temps que la présence du même phénomène en slave (*mrŭtvŭ* d'après *živŭ*) n'est pas significative, parce qu'elle serait indépendante de l'innovation latino-vénète. Mis à part le fait que l'on a un phénomène analogue aussi en celtique (vieil irl. *marb*, gall. *marw* « mort » d'après *beu* et *biw* « vivant »), il est tout à fait normal qu'un membre d'un couple antithétique soit influencé, même phonétiquement, par l'autre membre ; ainsi le couple *gravis* : *levis* du latin classique fut remplacé en latin pré-romain par **grevis* : *levis*. Des faits de ce genre peuvent arriver à n'importe quel moment et ne prouvent rien d'un point de vue dialectologique.

Dans d'autres cas, enfin, on a l'impression que certaines coïncidences sont complètement dues au hasard. La conjonction vénète *ke* peut être sans aucun doute comparée au lyc. *se* « et ». Mais cela ne nous autorise pas à en déduire un indo-européen **ḱe* « et », ni même à y voir une innovation commune au lycien et au vénète. Dans les deux cas, nous avons des spécialisations identiques, mais indépendantes, d'un élément **ḱe*, dont on ne peut rien dire avec certitude, si ce n'est qu'il ne représente pas une ancienne conjonction.

Ces quelques éléments significatifs d'un point de vue dialectologique permettent donc de mettre en relief deux faits. Tout d'abord, la présence de différents éléments archaïques, qui ne peuvent avoir de termes de comparaison dans la zone occidentale que s'ils en ont également dans la zone orientale. L'un des exemples typiques est le substantif *donom* « don » : ici, ce n'est pas la coïncidence avec le lat. *dōnum*, o. DUNUM, vieil irl. *dán* qui est significative (contre gr. δῶρον, vieux sl. *darŭ*, arm. *tur*), mais celle avec le véd. *dāna-* qui garantit le caractère archaïque du lexème.

Le prétérit *doto* « donna », qu'il faut certainement interpréter comme /*dōto*/, c'est-à-dire comme un prétérit athématique moyen avec vocalisme analogique de l'actif, est également un archaïsme; à sa base il y a un **dăto* immédiatement comparable au véd. *(a)dita* (issus tous les deux de *dH_3-*to*). Les autres langues occidentales ne présentent rien de semblable.

À côté de *doto* on a *donasto*, aoriste du verbe dénominatif issu de *donom*; et l'archaïsme de la formation athématique est prouvé par la comparaison avec le véd. *(a)stosta* « loua » et non pas avec le vieil irl. *-car* « aima » (**kara-s-t*). Cette catégorie verbale n'existe pas dans les dialectes italiques, et elle a été soumise à de fortes altérations en grec et en latin; le vénète la conserve très fidèlement.

Toujours dans le domaine du verbe, le vénète conserve le participe moyen en **-mno-/-meno-* (*karanmns* « Aimé », *alkomno* « renforcés »). D'autres formes verbales ont des interprétations plus controversées; toutefois, l'archaïsme de l'élément *-r* dans *toler*, *(d)idor* est évident.

À côté de la conservation de plusieurs éléments archaïques, on peut observer deux points très liés entre eux : l'absence presque totale d'éléments qui l'unissent aux langues indo-européennes centrales, et un nombre significatif d'éléments qui l'unissent aux langues occidentales.

En ce qui concerne le premier point, on ne peut pas accorder d'importance au rapprochement entre *es* (**ens*) « jusqu'à » et le grec εἰς puisque cette forme grecque est une innovation récente et que de nombreux dialectes grecs conservent ἐν + acc. Le rapprochement entre le nom *voltiio* et le dor. λῆν est assez incertain; celui avec le lat. *velle* serait beaucoup plus évident. Tout se réduit, en fait, à l'isoglosse lexicale entre le toponyme *Tergeste* (et *Opitergium*) et le vieux sl. *trŭgŭ* « marché »; mais rien ne nous garantit que ces deux toponymes signifient réellement « marché ».

En ce qui concerne le deuxième point, les matériaux sont nombreux. Même si l'on ne veut pas attribuer une attention particulière au traitement des sonores aspirées indo-européennes – qui, à notre avis, a eu certainement lieu de façon indépendante dans chaque langue – on remarquera le gén. sing. en *-i* (*keutini* « de Keutinos ») que l'on ne retrouve comme tel qu'en latin, falisque, messapien et celtique. De la même manière, le dat. pl. en *-bos* (*louderebos* « aux enfants ») ne trouve de correspondant qu'en latin et dans certains dialectes du celtique continental; on ne peut pas établir si

l'o. *-fs*, *-ss*, ombr. *-s* remontent à **-bhos* ou à **bhis*. La désinence **-bhos* représente donc une isoglosse typiquement occidentale, même si l'élément *-bh-* se retrouve aussi en véd. *-bhis*, *-bhyas*, *-bhyām*.

La racine indo-européenne *$dheH_1$-/dhH_1- apparaît avec un élément *-k-* dans le prétérit *hvagsto* « fit », comme cela est également le cas en latin (*faciō*, *fēcī*) et en italique (o. *fakiiad*, ombr. *façia*) ; le rapprochement avec le grec ἔθηκα resterait très partiel. L'adverbe vénète *eik* « ici » ne peut être comparé qu'au lat. *hic* (*heice*) et au falisque *fe/hec/he* ; le pronom réfléchi *sselboisselboi* « à soi-même » ne peut l'être qu'avec le v.h.a. *der selbselbo*.

Même si le lexique a peu de valeur probante dans le domaine de la dialectologie, à cause de la facilité avec laquelle un lexème peut être éliminé ou emprunté, il est toutefois significatif que l'analyse du lexique vénète s'accorde tout à fait avec ce que l'on a observé jusqu'à présent.

Les éléments qui sont très certainement archaïques et qui ont une distribution significative sont peu nombreux. Mis à part le *donom* déjà cité, nous pouvons rappeler **toutā* « peuple », attesté de l'Irlande jusqu'en Iran ; *augar*, qui désigne une offre, présuppose une valeur de « force, renforcement » que l'on retrouve en véd. *ojas-* et en av. *aojah-*.

Les éléments occidentaux, au contraire, sont très nombreux et concernent des secteurs importants de la vie sociale. Nous rappellerons ici le nom de l'hôte (vén. *hosti-*, lat. *hostis*, got. *gasts* ; le sl. *gostĭ* semble être un emprunt au germanique) ; le nom du fils (vén. *fīlia*, lat. *fīlius*, fal. *hileo*) ; le nom des enfants légitimes (vén. *louderobos*, lat. *līberī*) ; le nom des dieux (vén. *aisun*, dial. ital. *aisos. esos*) ; la racine de sens « aimer » (vén. *karanmns* « Aimé », lat. *carus*, v. irl. *caraimm* « j'aime », *carae* « ami », all. *Hure*) ; le nom du soldat (vén. *miles*, lat. *miles*, s'il ne s'agit pas d'un emprunt au latin).

Ces deux phénomènes – peu d'éléments archaïques, richesse d'éléments typiquement occidentaux (même si le concept de richesse doit, dans ce cas, être rapproché de la pauvreté des matériaux) – laissent à penser que le vénète a acquis son *facies* dans la période historique. Il se trouvait ainsi en contact avec la zone celtique et la zone germanique ; les Latins et les Italiques, en marche vers leurs habitats définitifs, ont dû passer par là ; et c'est là que, peut-être, des contacts avec le monde italique avant l'invasion gauloise furent possibles.

6

Les langues dont nous avons parlé jusqu'à présent sont, dans l'Italie antique, les seules parmi les *Restsprachen* qu'on puisse, à notre avis, tenter d'insérer dans la dialectologie indo-européenne. Ce n'est pas le cas pour les autres langues, car l'interprétation linguistique des matériaux épigraphiques apparaît trop souvent mal fondée ou même totalement erronée. Un cas typique nous est offert par le mess. *klaohi*, attesté dans le syntagme *klaohi zis*, qui, dans l'exégèse traditionnelle, était interprété comme « écoute, Zeus » et était comparé au véd. *śroṣi* « écoute » ; ils proviendraient tous les deux d'indo-européen _*ḱleu-si_. L'impératif messapien, qui ne peut être rapproché que du védique, apparaissait comme un élément de vénérable archaïsme.

Une première critique de cette analyse fut donnée par Schmid (1962), qui observa que *zis*, comme forme sigmatique, ne peut pas être un vocatif et que, par conséquent, *klaohi* ne peut pas être la deuxième personne d'un impératif. Toutefois, cette critique n'était pas décisive puisqu'ailleurs aussi, les exemples de vocatifs sigmatiques ne manquent pas (véd. *somapāḥ*, gr. ἄναξ).

Mais le coup décisif à ce rapprochement fut porté par Szemerényi (1966), qui démontra que les prétendus impératifs védiques en *-si* sont en réalité des formes haplologiques du subjonctif de l'aoriste sigmatique : *yakṣi* « sacrifie » < _*yakṣasi_, etc. Notre *śroṣi* (qui est un hapax) ne remonte pas directement à cette formation (l'aoriste de _*ḱleu-_ est athématique et non pas sigmatique), c'est une création analogique occasionnelle sur d'autres « impératifs » en *-si*. Toute possibilité de rapprochement entre la forme védique et la forme messapienne n'est donc plus, à ce point, possible. Mais est-ce que cette dernière forme peut être comparée à l'optatif tokh. B *klyausi* « écoute » ? La ressemblance, je le crains, n'est qu'apparente, puisque *klyausi* remonte à _*ḱlēus-oi-t_ (cf. skr. *śroṣet*), là où le messapien ne peut remonter à la même forme (cf. *berain*, venant probablement de _*bher-oi-nt_).

Dans d'autres cas, nous nous trouvons en face d'analyses qui peuvent paraître légitimes, mais qui demeurent incompatibles entre elles. Le mess. *pido* est généralement comparé au véd. (*api-*)*dāt* « donna » (_*epi-deH_$_3$_-t_), et le mess. *hipades* est analysé comme étant la continuation de _*dheH_$_1$_-s-t_ (aoriste sigmatique de _*dheH_$_1$- « poser, dédier »). Chacune de ces interprétations est possible, mais le fait que deux racines symétriques évoluent différemment, l'une en conservant l'ancien aoriste athématique, l'autre en présentant un aoriste sigmatique, n'est pas sans mettre dans l'embarras. De Simone (1988) avait déjà remarqué cette difficulté structurale

et il la résolut en admettant que les deux formes remontent à deux plans synchroniques différents. Mais de cette manière, le problème est déplacé, et non pas résolu ; on peut se demander en effet pourquoi **dō-t* n'est pas devenu **dō-s-t* quand **dēt* est devenu **dē-s-t*.

Nous devons donc être conscients que dans cette situation de profonde incertitude au niveau de l'exégèse, l'insertion du messapien dans le cadre de la dialectologie indo-européenne aurait peu de sens, tout comme il serait assez inutile de reporter dans le domaine de la dialectologie les témoignages anciens sur les rapports ethniques très étroits existant entre les Messapiens et les Illyriens, et de s'appuyer sur l'affinité entre l'onomastique messapienne et l'onomastique balkanique pour conforter cette théorie, qui ne peut avoir une valeur probative supérieure à la valeur de l'affinité entre l'onomastique latine et l'onomastique étrusque. On ne doit pas confondre le voisinage culturel avec le voisinage linguistique.

Des réserves identiques et encore plus fondées devront être faites à propos du sicule, que les chercheurs ont apparenté parfois au latin, parfois aux dialectes italiques, parfois avec d'autres *Restsprachen* comme le vénète, le messapien ou l'illyrien.

Le gaulois d'Italie, représenté par un faible nombre d'inscriptions, ne peut pas être séparé du gaulois de Gaule et ne peut donc avoir de traitement autonome ; tous deux rentrent dans le cadre plus général du celtique. On admet généralement aujourd'hui que le lépontique est un dialecte celtique ; mais il est à craindre que cela n'ait été établi sur la base d'interprétations et de reconstructions qui ne sont pas tout à fait convaincantes.

Bibliographie

(arrêtée en 1990)

BEELER M.S. (1949). – *The Venetic Language*, Berkeley & Los Angeles.

– (1966). – «The Interrelationships within Italic», *Ancient Indo-European Dialects*, ed. by H. Birnbaum and J. Puvhel, Berkeley & Los Angeles, p. 51 *sq*.

BENVENISTE E. (1935). – *Les Infinitifs avestiques*, Paris.

CAMPANILE E. (1985). – «Riflessioni sui più antichi testi epigrafici latini», *AION* 7, p. 89 *sq*.

DE SIMONE C. (1988). – «Iscrizioni messapiche della grotta della poesia», *ASNP* 18, p. 325 *sq*.

DEVOTO G. (1929). – «Italo-greco e italo-celtico», *Silloge Ascoli*, Torino, p. 200 *sq* (= *Scritti minori*, vol. I, Firenze, 1958, p. 129 *sq.*

LEJEUNE M. (1977). – *Manuel de la langue vénète*, Heidelberg.

LEUMANN M. (1977). – *Lateinische Laut- und Formenlehre* (5. Auflage), München.

MARINETTI A. (1985). – *Le iscrizioni sudpicene*, vol. I, Firenze.

MARSTRANDER C. (1929). – «De l'unité italo-celtique», *NTS* 3, p. 241 *sq.*

PFISTER R. (1977). – = F. Sommer, *Handbuch der lateinischen Laut- und Formenlehre*, 4. Auflage, I. Band, von R.P., Heidelberg.

PORZIG W. (1960). – «Altitalische Sprachgeographie», in *Festschrift Krause*, Heidelberg.

RIX H. (1957). – «Sabini, Sabelli, Samnium», *BNF* 8, p. 127 *sq.*

SCHMID W. (1962). – Messapisch *klaohi, klohi*, *Studi Salentini* 14, p. 332 *sq.*

SZEMERÉNYI O. (1966). – «The origin of the Vedic «Imperatives» in *-si*», *Lang.* 42, p.1 *sq.*

VENDRYES J. (1934). – «La mort des langues» Conf. Inst. Ling. 1, p. 1 *sq.* (= *Choix d'études linguistiques et celtiques*, Paris, 1952, p. 39 *sq.*).

VETTER E. (1962). – «War das Sikulische eine italische Sprache?», *Glotta* 40, p. 62 *sq.*

CHAPITRE 15

LE CELTIQUE

Enrico CAMPANILE

1

La définition du celtique dans le cadre de la dialectologie indo-européenne ne peut se limiter aujourd'hui à la reconstruction de la position que ces dialectes auraient occupée dans l'espace au moment de la dissolution de l'unité indo-européenne, ni à l'explication des articulations généalogiques entre la langue mère et les parlers celtiques historiquement attestés. C'est plutôt une opération d'ordre historico-culturel (et, comme telle, toujours provisoire et subjective), où des éléments archaïques et des éléments novateurs sont utilisés comme preuves, les uns de la conservation d'une tradition linguistique originelle, les autres de contacts et d'expériences nouvelles que la culture celtique élabora en commun avec d'autres cultures ou qu'elle vécut de manière autonome.

De ce point de vue, la vieille question de l'unité italo-celtique apparaît non seulement mal fondée d'un point de vue technique, mais, pis encore, peu importante dans la mesure où, quand bien même on pourrait la démontrer, elle resterait toujours un épisode particulier d'une histoire bien plus vaste et complexe.

Le concept de définition dialectale en tant que résumé de la préhistoire d'une langue, que nous proposons ici, est déjà clairement implicite dans diverses enquêtes qui utilisent le matériau celtique non pas à des fins classificatoires et généalogiques, mais pour en établir l'ancienneté, la chronologie relative et les rapports ponctuels avec des matériaux analogues d'autres langues indo-européennes; de ce point de vue, l'article dans lequel Schmidt (1986) examine les rapports du celtique avec les autres langues anciennes de l'Europe est aujourd'hui encore exemplaire.

Mais la difficulté majeure d'une telle enquête naît du fait qu'il y a aujourd'hui parmi par les chercheurs de très fortes divergences dans l'analyse et dans la reconstruction de presque tous les secteurs cruciaux de la morphologie et de la syntaxe celtiques. L'interprétation des différents faits grammaticaux que l'on offre dans ces pages ne se fondera donc que rarement sur le consensus des chercheurs, comme on pourrait le souhaiter; ce sera plus souvent celle qui semble la plus probable à l'auteur de ces lignes, bien qu'il ne lui soit pas possible d'en donner ici une justification appropriée; de même, le fait que des interprétations différentes ne soient pas évoquées n'implique pas forcément que celles-ci soient méconnues de lui.

2

La volonté de n'utiliser que les matériaux les plus anciens parmi tous ceux qui sont offerts par chaque langue, est présente aujourd'hui chez beaucoup de chercheurs dans la reconstruction de l'indo-européen – et, plus généralement de tous les phases linguistiques antérieures aux documents – : ainsi, par exemple, on ne prend pas en considération l'indien ancien, mais uniquement le védique; et non pas tout le védique, mais seulement celui du Rig-Veda; mieux encore : non pas tout le Rig-Veda, mais seulement les hymnes de certains livres bien précis. Cette rigueur documentaire entraîne, en réalité, un renversement méthodologique : l'archaïsme linguistique est fonction non plus de la possibilité de comparaison, mais de la pure chronologie textuelle. À notre avis, on peut objecter à cette façon de procéder que, s'il est statistiquement probable qu'on trouve plus d'éléments archaïques dans la phase ancienne d'une langue que dans une phase récente, ce fait est sans importance sur le plan heuristique , dans la mesure où l'on ne peut pas pas en tirer d'appréciations spécifiques pour les cas particuliers.

Cette conclusion est importante en particulier pour l'appréciation du celtique continental, qui, d'un côté, représente de loin la phase la plus ancienne du celtique, mais qui, de l'autre, ne peut être utilisé que dans une petite mesure pour la reconstruction à cause de la rareté des matériaux et de l'extrême incertitude qui pèse souvent sur leur interprétation. Ne nous laissons pas, en effet, induire en erreur par le grand nombre de traductions qui accompagnent les éditions des textes celtiques continentaux : dans la majorité des cas, nous nous trouvons malheureusement devant des hypothèses très douteuses ou même devant des erreurs manifestes. D'où la nécessité de fonder l'analyse dialectologique presque exclusivement sur le matériau insulaire; mais cela ne bouleversera cependant pas, à notre avis, les résultats de l'enquête; et cela pour deux raisons.

La première est d'ordre factuel. Il est vrai, en effet, que le celtique continental peut conserver des éléments archaïques inconnus du celtique insulaire : la conjonction celtibère *uta* en est un exemple évident, avec des correspondants en indo-iranien, mais non en gaëlique ou en britonnique; mais il est vrai aussi que dans d'autres cas, la phase la plus archaïque peut être celle du celtique insulaire : le vieil-irl. *imté* «protège» (qui continue un injonctif) est, sur le plan de la typologie indo-européenne, plus archaïque que la forme correspondante celtibère *amPiTiseTi* (qui continue un subjonctif).

Le seconde raison est d'ordre méthodologique. L'histoire linguistique qui est reconstruite ici est une histoire de longue durée : les catégories examinées ont une vie séculaire ou millénaire, même quand elles sont clairement novatrices, comme c'est le cas pour le futur irlandais en *-f-*, qui apparaît déjà aux débuts de la tradition écrite, et qui existe aujourd'hui encore. Dans un cadre de ce genre, l'absence d'un témoin peut appauvrir la reconstruction, mais elle n'altère pas les grandes lignes.

3

Dans la définition dialectologique des parlers celtiques, les phénomènes phonétiques sont peu révélateurs; il s'agit en effet soit de phénomènes récents qui ne recouvrent même pas toute l'aire celtique (comme c'est le cas pour la lénition, qui ne concerne que le celtique insulaire), soit de phénomènes antérieurs aux plus anciens documents mais monoglottes (comme la restructuration du système consonantique), soit encore de phénomènes banals sur le plan typologique et non utilisables à des fins classificatoires.

En réalité, la seule donnée phonétique qui pourrait être de poids d'un point de vue classificatoire serait l'assimilation de **p*... *k^w en *k^w... *k^w, si elle avait vraiment le caractère d'une loi phonétique valable pour le celtique et pour le latin (tout en demeurant cependant invérifiable en ce qui concerne les dialectes italiques). Mais en fait, cette loi présumée telle – qui constituait autrefois l'un des piliers de l'unité italo-celtique – se fonde seulement sur deux exemples : lat. *coquō*, gall. *pobi* « cuire » (*pek^w-) et lat. *quinque*, vieil-irl. *cóic*; gall. *pymp* « 5 » (*penk^{w}e), auxquels on oppose deux exemples qui vont dans le sens contraire : gaul. *(H)ercynia (silva)*, mais lat. *quercus*, tous deux à partir de *perk^w-, et vieil-irl. *déec* « 10 » (*dwi-penk^{w}om). Il ne s'agit donc pas d'une loi phonétique que l'on peut faire remonter à une phase unitaire des deux groupes linguistiques, mais seulement d'un phénomène occasionnel d'assimilation.

4

On peut considérer que la caractéristique la plus importante des parlers celtiques réside dans leur archaïsme, caché sous un voile d'innovations manifestes, en particulier sur le plan phonétique. Le comparatiste éclaire cet archaïsme en utilisant un double instrument. D'un côté il identifie des éléments qui disposent d'un terme de comparaison dans l'aire orientale du monde indo-européen (surtout en indo-iranien), mais non dans l'aire centrale : la séparation géographique de ces éléments garantit leur appartenance au patrimoine le plus ancien (l'affirmation contraire, en effet, exigerait l'hypothèse improbable d'innovations identiques mais indépendantes). De l'autre côté, le comparatiste met en évidence des éléments que l'on retrouve aussi en hittite (et éventuellement dans l'aire orientale, mais non dans l'aire centrale) et il y trouve une confirmation du caractère indo-européen de l'élément hittite soumis à comparaison (en ce sens qu'il a un correspondant en celtique), ainsi que le caractère archaïque de l'élément celtique (en ce sens que le hittite est riche d'éléments non indo-européens, mais que ceux qui sont indo-européens présentent une forte empreinte archaïque).

La caractéristique constante de cette comparaison est qu'en principe, l'élément celtique n'est pas immédiatement analysable mais le devient seulement après que l'on a identifié le terme correct pour la comparaison. Nous voulons dire, par exemple, qu'une forme verbale comme v.irl. *-té* ne manifeste plus, en raison de toute une série d'innovations phonétiques, aucun trait qui révèle son ancienne nature d'injonctif de l'aoriste sigmatique;

celle-ci apparaît uniquement à travers la comparaison avec des formations indo-iraniennes que l'on analyse sans difficulté. De la même manière, la fonction originelle de l'élément *no-* est démeurée obscure jusqu'à ce que sa valeur primitive de *satzeinleitende Partikel* ait été dévoilée grâce à la comparaison avec le hittite *nu-*.

Il ne sera pas exagéré de considérer que cet archaïsme n'est rien d'autre que l'aspect linguistique de ce conservatisme et de cet isolement qui, plus généralement, caractérisent la culture celtique dans ses aspects les plus divers.

Nous allons examiner à présent une série d'éléments linguistiquement archaïques qui figurent en celtique.

5

Le caractère archaïque de domaines particuliers du lexique celtique (comme la religion, le droit, la guerre) fut déjà reconnu par Vendryes (1917), qui identifia divers éléments que ce lexique partage avec l'indo-iranien, mais non pas avec les langues de l'aire centrale. D'autres lexèmes, sûrement très anciens, ont été mis en lumière par la suite ; nous en rappelons ici quelques-uns :
v.-irl. *airech* « concubine » : av. *pairikā* (Thurneysen, 1924) ;
v.-irl. *sétig* « femme » : anc. ind. *satī̄* (Campanile, 1966) ;
gaul. *mediosamonios*, nom d'un jour : av. *maiδiiōi.šəma-*, nom du dieu de la seconde partie de l'année (cf. Kellens, 1974, 399) ;
gall. *herw* « raid » : hitt. *saru* « butin » (Watkins, 1976) ;
gall. *chwŷf* « vibration » : av. *xšuuiβ-* (Campanile, 1980) ;
v.-irl. *-bria* « qu'il nuise » : véd. *bhrīṇánti* « ils frappent » (Wagner, 1982) ;
gaul. *bar(-dos)* « (faiseur de) louange » : véd. *gir-* « louange » (Campanile, 1980a) ;
gall. *traidd* « pénétration » : véd. *tṛṇátti* « il troue » (Campanile, 1991).

Un autre archaïsme lexical facilement observable, à la fois lexical et morphologique, est représenté par le nom de la femme : vieil-irl. *bé* (neutre). Celui-ci remonte à *$*g^w en$* qui, en raison de son genre grammatical neutre et de l'absence de suffixe, se révèle plus archaïque que le véd. *gnā-* « femme divine » ($*g^w neH$- thème consonantique, et non en *-ā* comme dans Wackernagel 1975, § 56b). Le gén. *bé* continue l'indo-européen $*g^w en$-*s*, identique dans sa structure à **dem-s* « de la maison » (véd. *dan*, av. *dǝ̄ng*). Aucune autre langue indo-européenne ne présente ce nom comme un neutre et sans suffixe (Campanile, 1977, Hamp, 1979).

6

Sur le plan de la morphologie nominale, l'archaïsme le plus remarquable est représenté par la conservation des formes spécifiques du féminin pour les numéraux « 3 » et « 4 » ; on retrouve le même phénomène en indo-iranien et en tokharien :

« 3 » : v.-irl. *teoir*, gall. *tair*, bret. *teir* (m. : v.-irl., gall. bret. *tri*) : véd. *tisráḥ* (m. *tráyaḥ*), av. *tišrō* (m. *ϑrayō*), tokh. B *tarya* (m. *trai*). L'existence d'un anc. nord. *þrijor* (fém.) est niée par Marstrander (1930, 316) ;
« 4 » : v.-irl. *cetheoir*, gall. *pedair*, bret. *peder* (m. gall., : v.-irl. *cethir*, gall. *pedwar*, bret. *pevar*) : véd. *cátasraḥ* (m. *catvā́raḥ*), av. *cataŋrō* (m. *caϑvarō*), tokh. B *štwāra* (m. gall., *štwer*). Le gaul. *tidrus* (ou *tidres*), dans lequel on voyait autrefois le fém. « 4 », est lu aujourd'hui *tibrus*, avec un sens inconnu (Ellis Evans, 1977).

7

Il semblerait qu'en celtique l'expression du nombre ait conservé la trace d'une situation très archaïque, où la distinction entre couple naturel et couple occasionnel trouvait son expression aussi le plan formel ; il s'agit, en tout cas, d'un problème que l'on n'a pas encore suffisamment étudié ; il convient, par conséquent, de se limiter à le signaler.

Le celtique emploie le duel aussi bien pour le couple naturel (v.irl. *di lám* « les (deux) mains ») que pour le couple occasionnel (v.irl. *di mnaí* « deux femmes »), mais pour désigner un seul membre du couple naturel il utilise un composé avec *leth-* « moitié » (v.irl. *lethlám* « une main »). Or, une distinction morphologique du même type se retrouve en tokharien (Krause-Thomas, 1960, 76), où l'on oppose le paral (pour le couple naturel : *A ašäṃ*, B *ešane* « les deux yeux ») au duel (pour le couple occasionnel : B *pwāri* « les deux feux ») ; et le paral du tokharien, en outre, exprime également le duel elliptique, dont l'origine est certainement indo-européenne et qui, par l'intermédiaire du duel d'un membre, désigne le couple hétérogène, mais naturel du point de vue historique : tok. B *ñaktene* « le dieu et sa femme » (la paire à partir de *ñakte* « dieu »), exactement comme dans le véd. *Mitrā* « Mitra et Varuṇa » et dans le grec Αἴαντε « Ajax et son frère Teucros ».

Il semblerait donc qu'il y eût une structure cohérente et très archaïque qui donnait une expression morphologique aux contenus sémantiques (Campanile, 1980a) : si c'était le cas, l'opinion de Vendryes, qui voyait dans la situation du celtique le produit d'innovations tardives, serait désavouée.

8

Le celtique participe dans une large mesure à l'innovation qui amena à la disparition des présents athématiques au profit d'autres formations : tendance très ancienne comme le prouve la comparaison de hitt. *huiszi* « (il) vit » avec véd. *vásati* « (il) habite ». Mais il est probable qu'outre la copule **esti* (vieil irl. *is*, gall. *ys*), qui tend à demeurer partout athématique, un autre présent radical survit en celtique.

Dans le paradigme irlandais de « aller » on voit converger des formes étymologiquement différentes; « il va » est *téit*, que Bergin (1938) fit remonter à **ten-ti* (pour l'emploi intransitif de « tendre » dans le sens de « mouvoir en direction de », cf. lat. *quo tendis* ?, gr. ποῦ τείνεις ;) étymologie qui, malgré les réserves de Meid (1972) et d'autres chercheurs, apparaît convaincante. Mais le fait que le védique forme un aoriste (*átan* « il tendit ») et non pas un présent à partir de la racine **ten-* était déjà considéré par Bergin comme un obstacle à son hypothèse. Cette difficulté peut être surmontée à l'aide de la comparaison avec le grec τένων « tendon (ce qui tend le muscle) », qu'il n'y a pas lieu de considérer comme un participe aoriste (Strunk, 1967). La conservation de cette forme athématique isolée dépend probablement du fait que le celtique avait englobé **tenti* dans le paradigme de « aller » et qu'il avait perdu les autres formes du paradigme avant que de nouvelles formations de présent à suffixe, comme **ten-ye-* ou **tn̥-n-ew-*, aient transféré **(e-)ten-t* de l'imparfait (ou injonctif du présent) à l'aoriste.

9

En indo-européen, il existait une série de particules sans aucun contenu sémantique autonome, et qui avaient pour fonction d'introduire la phrase (*satzeinleitende Partikeln*) et de servir de support à d'éventuels éléments atones, situés en position enclitique selon la loi de Wackernagel (1892). Ces particules, qui en hittite et dans les autres langues anatoliennes sont d'une attestation immédiate (Carruba, 1969), ont été progressivement identifiées aussi en celtique (Watkins, 1963), avec des fonctions encore extrêmement archaïques, c'est-à-dire à un niveau différent des autres langues indo-européennes où, quand elles subsistent, elles apparaissent plus ou moins transformées ou fossilisées.

Nous avons donc en vieil-irlandais le préfixe asémantique *no-* = hitt. *nu* : *no-m-bered* « il me portait » (hitt. *nu-mu*), *na-n-oirg* « il le tue » (hitt. *nan-*), *noch* (hitt. *nukku*) à valeur affirmative (*noch ém am israhélde* « en effet je suis hébreu »); le *to* dit « vide » (Dillon, 1963 : « vide » en tant qu'asémantique et distinct donc de la préposition et préfixe *to* « vers, à ») = hitt. *ta* : *do-chlaid* « qui creuse » (**to-yo-*); la conjonction *sech* « et en vérité » (hitt. *sukku*; le vocalisme irlandais peut être analogique de *sech* « plutôt, autrement », cf. lat. *secus*). A la lumière de ces faits, on peut aussi se demander si le v.-irl. *co n-* (en début de phrase et avec sa pure valeur de conjonction) n'était pas lui aussi une particule introductive comparable au hitt. *kan(a)* et différente donc du *co n-* consécutif.

En indo-européen il existait également des modificateurs verbaux qui, sur le plan syntaxique, pouvaient se comporter comme des particules introductives de phrases et avoir aussi fonction de support d'éléments enclitiques successifs. Dans les deux cas, la structure de la phrase admettait le verbe en position finale :

véd. : *abhí no devī́r ávasa mahā́ḥ s'ármaṇā nṛpátnīḥ áchinnapatrāḥ sacantāṃ* « Die Göttinnen sollen mit ihrer Gnade, mit ihrem größen Schirm uns zu Seite stehen, die Herrscherfrauen mit ungebrochenen Fittichen » (R̥V 1, 22, 11); av. *ā mōi rafəδrāi zavə̄ng jasatā* « répondez à mon invocation d'aide » (Y. 28, 3); hitt. *n-as-san ANA*^{GIŠ}BANŠUR.AD.KID.*ANA PANĪ* DINGIR^{LIM}*tehhi* « et moi de la claie je les pose sur la table devant la déesse » (Carruba, *op. cit.*, 14). La même structure se conserve dans des textes poétiques irlandais : *nom Choimmdiu coíma* « le Seigneur m'assiste ».

Dans le cadre de choix qui ont fini par ne plus être analysables, il était néanmoins possible que le verbe ait été placé avec les autres enclitiques : véd. : *ā́ tvā vahantu hárayo vṛ́ṣaṇaṃ sómapītaye* « (que) les chevaux te portent, toi, le taureau, pour boire le soma » (R̥V 1, 16, 1);

gr. πρό μ'ἔπεμψεν ἄναξ ἀνδρῶν 'Αγαμέμνων «Agamemnon, protecteur de son peuple, m'a envoyé» (A 442) (plus anciennement *πρό μ'ἔπεμψεν). C'était là la séquence normale dans les dialectes insulaires, quand le verbe est précédé d'un quelconque préfixe : vieil irl. *atut-áigfedar* «chacun te haïra», v. gall. *namercit mi nep leguenid* «que nul ne me demande de l'allégresse» (*na-m-ercit*). En l'absence d'éléments «préfixaux» (*satzeinleitende Partikeln* et modificateurs verbaux), le verbe indo-européen pouvait occuper la position initiale en devenant tonique et en acceptant d'éventuels éléments enclitiques :

véd. *gā́yanti tvā gāyatríṇaḥ* «les chantres te chantent, toi» (R̥V 1, 10, 1)
av. : *dāidī tū ārmaitē vīštāspāi īšəm* «accorde donc, ô Pitié, à Vistaspa la force» (Y. 28, 7)
gr. φαίνεταί μοι κῆνος ἴσος θέοισιν «celui-là me paraît égal aux dieux» (Sappho 199 P.)

C'est là le type normal dans la prose irlandaise archaïque : *cechardut cách* «chacun t'aimera», qui plus tard laissera la place à la généralisation du type *no-t-chechra* (particule introductive + pronom + verbe) que nous avons examinée auparavant.

Dans ce dernier syntagme, l'accent tombe sur le verbe et cela constitue une innovation par rapport au modèle indo-européen qui comportait une accentuation du préfixe. Mais la situation originelle se conserve avec la loi de Bergin (1938) : quand le verbe préfixé est en fin de phrase, l'accent tombe sur le préfixe. Nous avons ainsi *maicne nAilb áirmi* «tu énumères la descendance d'Alb», là où dans une prose usuelle on aurait *adrími maicne nAilb*. Le type *maicne nAilb áirmi* est identique sur le plan syntaxique au véd. *devó devébhir ā́ gamat* «vienne le dieu avec les dieux» (R̥V 1, 1, 5) : préfixe tonique + verbe atone, tous deux en position finale.

Les nombreuses possibilités de position et d'accentuation du verbe celtique (Wagner, 1967) reflètent souvent, en réalité, une situation indo-européenne archaïque.

10

L'une des caractéristiques du celtique réside dans le fait que le verbe a deux paradigmes différents : un paradigme avec des formes «brèves» quand il a un préfixe, et un autre avec des formes «longues» quand il n'en a pas. Cette opposition, qui est constante en irlandais (*no-m-beir* «il me porte» : *beirid* «il porte») a dû exister autrefois aussi dans les dialectes

britonniques, comme le démontrent des vestiges tels que *trengid golud ni threing molud* « la richesse passe, la gloire ne passe pas » (la négation *ni* est un élément préfixal). On a avancé beaucoup d'hypothèses sur l'origine de cette opposition, hypothèses que l'on peut quand même regrouper autour de deux courants principaux. D'après le premier, les deux types sont issus d'une même base de départ (qu'il s'agisse des désinences primaires ou des désinences secondaires de l'indo-européen) et la différenciation se serait faite dans un cas par la perte, dans l'autre par l'adjonction d'un élément terminal. Pedersen, par exemple, partait de formes verbales avec des désinences secondaires et faisait remonter la forme citée *beirid* à **bheret* (injonctif) + *is* (pronom), là où *-beir* continuerait l'injonctif simple. D'après l'autre interprétation, à laquelle nous adhérons, la différence reflète en substance l'ancienne opposition entre désinences primaire (indicatif) et désinences secondaires (injonctif) ; d'où : *beirid* = véd. *bharati*, *-beir* = véd. *bharat*. Le fait que des formes d'injonctif aient survécu de façon autonome en celtique (c'est-à-dire même dans des formations verbales sans préfixe), est mis en évidence par des imparfaits comme *nascad* « liait », à la base duquel il faut voir sans aucun doute un injonctif moyen en *-to*.

Mais les raisons de la distribution de l'indicatif et de l'injonctif sont peu claires en celtique. On pourrait penser que les formes préfixées généralisent l'injonctif en ce sens qu'elles comportent des valeurs modales ou temporelles que l'on ne retrouve pas dans les formes correspondantes non préfixées, mais qui concordent avec les anciennes fonctions de l'injonctif indo-européen. Il n'en demeure pas moins que l'injonctif n'est pas moins présent en celtique qu'il ne l'est en indo-iranien.

11

L'irlandais ancien possède un subjonctif sigmatique (liste complète dans Watkins 1962, 128 et *sq.*), composé du thème verbal à degré plein + *-s-* ; la formation est en partie athématique (*géis* « qu'il prie » : *g^whedh-s-ti*), en partie thématique (*-gessam* « que nous priions » : *g^whedh-s-o-me*). On en a des vestiges aussi en britonnique (*duch* « qu'il mène » : *deuk-s-* ; *gwares* « qu'il aide » : **upo-ret-s-*, etc.) et peut-être aussi en celtibère (*amPiTiseTi* « qu'il protège », d'après Schmidt, 1976 à partir de **steigh-s-e-t(i)*).

Cette formation est issue directement de l'aoriste sigmatique indo-européen : les formes athématiques continuent l'injonctif (véd. *bhār* : **bher-s-t*, av. *dārəšt* : **dher-s-t*), les formes thématiques continuent le

subjonctif (véd. *bhakṣat*, av. *baxšaitī* «(qu'il) attribue», gr. ἐρύσσομεν). Comme preuve d'archaïsme, on pourra remarquer qu'en celtique ces formations se sont conservées sans innovations, là où en indo-iranien, elles ont allongé la voyelle du radical (le caractère récent de cet allongement a été démontré par Watkins 1966) et en grec le type ἐρύσσαμεν a remplacé, dès l'époque prélittéraire, le *Ϝερυσμεν originel.

C'est de l'injonctif de l'aoriste sigmatique que provient également une petite série d'impératifs irlandais, comme *at-ré* «surgis» et quelques autres, qui ne sont pas du tout des emplois spécialisés du subjonctif (sigmatique) irlandais, parce qu'à la deuxième personne du singulier celui-ci est thématique (*-reiss* «que tu surgisses» **reǵ-s-es*). Cela signifie que *at-ré* provient directement de l'injonctif indo-européen; sa valeurd'impératif caractérise également l'injonctif védique (*abhí kranda stanáya gárbham ā́ dhāḥ* »mugis, tonne, dépose la semence», R̥V 5. 83. 7, Hoffmann, 1967, 261).

12

En irlandais, le futur sigmatique est lui aussi étroitement lié au subjonctif sigmatique, et il ne s'en différencie que par le redoublement : subj. *geis* «qu'il prie», fut. *gigis* «il priera» (*g^{w}*hedh-s-ti* : *g^{w}*hi-*g^{w}*hedh-s-ti*). Il y a cependant quelques futurs sigmatiques qui sont dépourvus de redoublement (comme *seiss* «s'assiéra», *reiss* «courra», *-lé* «sera couché») qui coïncident avec les subjonctifs correspondants (**sed-s-ti*, **ret-s-ti*, **legh-s-t*), ce qui laisse supposer qu'à l'origine le subjonctif et le futur sigmatique constituaient une seule formation, sans redoublement, – c'est-à-dire d'anciens subjonctifs et injonctifs de l'aoriste sigmatique – et que, au futur, le redoublement a été introduit par analogie avec le désidératif, qui en irlandais donnait un type plus récent de futur.

Au contraire, il n'y a pas de lien étroit entre le type *seiss* «s'assiéra» (**sed-s-ti*) et le futur italique (osque *emest* «interdira», ombrien *ferest* «portera») puisque ce dernier présente un morphème *-es-* par rapport au seul *-s-* de l'irlandais.

Un autre type de futur de l'irlandais est représenté par l'ancien désidératif. Celui-ci se conjugue comme le subjontif en *-ā-* et présente dans sa structure deux particularités qui ont un caractère manifestement innovateur, mais qui apparaissent aussi en indo-iranien, ce qui laisserait à penser que ce sont certes des innovations, mais très anciennes.

La première particularité est que le thème verbal peut se présenter avec un degré plein comme avec un degré réduit : vieil-irl. *-gignethar* « naîtra » comme véd. *jíjaniṣate* (**ǵi-ǵenH-*), mais v.-irl. *lilsit* « lècherons » comme skr. *lilikṣati* (**li-ligh-*). La seconde particularité est que la laryngale qui, dans les racines *seṭ* se trouvait en position présigmatique (véd. *jijñāsate* « il veut savoir » : **ǵi-ǵneH-*, av. *zixšnàŋhəmna-*), a été également étendue, par un processus de resegmentation, aux racines *aniṭ* (par ex. véd. *cíkīrṣati* « veut faire ») ; la comparaison du v. irl. *génaid* « blessera » avec le véd. *jíghāṃsati* semble donc loin d'être fortuite et suppose une forme plus ancienne **g^{w}hi-g^{w}hn̥-Hs-e-ti*. Cela explique aussi qu'en irlandais ce futur soit conjugué comme le subjonctif en *-ā-* (**R̥Hseti* > **Rāti*).

13

Que ce soit dans le système verbal celtique ou dans celui de l'italique, un élément *-r-* caractérise plusieurs types de formations :

a) verbes déponents : vieil-irl. *suidigidir* « suit », lat. *sequitur*, ombr. *terkantur* « inspectent » ;

b) formations passives impersonnelles (en tant que distinctes du déponent) : lat. *itur*, ombr. *ier* « on va », v.irl. *berair* « on porte », *carthair* « on aime », gall. *cerir* « on aime », *kenhittor* « on sonnera » ;

c) la troisième personne du pluriel du parfait : lat. *fuēre* (non pas *fuerunt*, qui vient de **-is-ont*), v.irl. *batar* « y furent ».

Ce parallélisme fonctionnel a été autrefois considéré comme une preuve évidente de l'unité italo-celtique (idée introduite dès 1861 par Lottner). Aujourd'hui, au contraire, nous savons que cet élément apparaît avec des fonctions analogues dans beaucoup d'autres langues indo-européennes (hittite, indo-iranien, tokharien, phrygien, arménien), si bien qu'en raison de sa diffusion et de sa distribution, il pourrait représenter plutôt un élément d'un archaïsme considérable. Il n'est cependant pas facile d'identifier dans deux ou plusieurs langues des morphèmes avec *r* qui se superposent parfaitement, d'un point de vue aussi bien phonétique que fonctionnel. Si nous considérons, par exemple, la troisième personne du pluriel du parfait, le matériau à notre disposition sera représenté par lat. *fuēre*, v.irl. *batar* (déjà cité), véd. *cakrúr* « firent », av. *àŋharə̄* « furent », tokh. A *prakär*, B *prekar* « demandèrent ». Malgré leur évidente affinité, aucune de ces formes ne coïncide de manière satisfaisante avec les autres ; d'un point de vue étymologique, en effet, nous avons :

lat. *fuēre* : **-ēre*,
v. irl. *batar* : **-nto + r* (*fitir* « sait » : **ri-*, si ancien pluriel),
véd. *cakrúr* : **-ṛs*,
av. *åŋharə̄* : **-V̇-r*,
tokh. A *prakär*, B *prekar* : **ṛ* (?), **ṛs* (?)

Tout se passe donc comme si l'on se trouvait devant un élément si archaïque que chaque langue l'a remodelé et réutilisé à sa façon, ce qui empêche de reconstruire sa forme et sa fonction originelles, si ce n'est au prix d'hypothèses arbitraires.

14

Un autre élément jouait traditionnellement en faveur de l'hypothèse italo-celtique, c'est le suffixe superlatif **is-m̥mo-* : lat. *pigerrimus* (**pigrisamos*), v. irl. *sinem*, v. gall. *hinham* « le plus vieux » (**senisamos*). Bien qu'une telle formation ne se retrouve pas par ailleurs (comme cela a été encore souligné par Cowgill 1970), sa valeur probatoire est toutefois très faible, puisqu'elle résulte du croisement de deux suffixes de superlatif, **is-to* et **-m̥mo-* (véd. *svā́diṣṭha-* « très doux », *madhyamá-* « médian »). La prétendue *lex Dybo*, selon laquelle la perte d'une laryngale prétonique se serait produite dans l'aire occidentale, n'a pas plus de valeur probante en faveur de l'unité italo-celtique, comme le voudrait Kortland (1981), car cette loi est contredite par des nombreux exemples et quand bien même elle existerait, elle devrait comprendre également le germanique et constituerait, dès lors, une preuve non pas en faveur, mais plutôt contre l'italo-celtique. Il suffirait, en effet, de considérer la distribution du lexème **wī̆ros* « homme » (lat. *vir*, v. irl. *fer*, gall. *gŵr*, got. *wair* en face du véd. *vīrá-*, lit. *výras* et surtout ombr. *veiro-*) pour comprendre que cette alternance, qu'il faut encore expliquer, n'a rien à voir avec l'unité italo-celtique.

On ne peut nier toutefois qu'il y ait des éléments communs au celtique, au latin et aux dialectes italiques, qui, même s'ils ne peuvent être utilisés à des fins de généalogie, suggèrent une phase de rapports intenses entre ces dialectes, si bien que, pour reprendre un terme employé par Vendryes, il est normal de voir entre eux une affinité particulière.

15

Le latin aussi bien que l'irlandais possèdent un subjonctif en **-ā-* qui était originellement indépendant du thème du présent : *advenat* en latin archaïque (face au plus récent *adveniat*), *-cria* « il achète » en irlandais (thème du présent *cren-*) ; dans les dialectes italiques, l'adaptation au thème du présent est déjà accomplie au début de nos documents : o. *fakiad*, ombr. *façia* « qu'il fasse ».

Ce cadre schématique a longtemps constitué l'un des fondements de l'unité italo-celtique, mais la situation est en réalité bien plus complexe. D'un côté, en effet, on a vu qu'un subjonctif en *-ā-* existe aussi en tokharien (Lane, 1959), de l'autre on a souligné les rapports possibles entre ce type de subjonctif et des formes de prétérit en *-ā-* attestées aussi dans d'autres langues indo-européennes, ce qui pourrait accroître d'une manière considérable l'aire de cette formation (Benveniste, 1951). De ce point de vue, l'observation de Schindler (1967), selon laquelle le thème verbal **kriy-ā-* en tokharien donne naissance au prétérit (*käryām* « nous achetâmes ») tandis qu'en irlandais il donne naissance à un subjonctif (*-criam* « que nous achetions »), pourrait avoir une valeur d'exemple. Un jugement objectif requiert toutefois que l'on considère au moins les faits suivants :

a) le latin, le celtique et les dialectes italiques ont une formation prétéritale en *-ā-* uniquement dans deux verbes : lat. *erat* = gall. *oedd* « était » (**es-ā-t*), et lat. *-bat* (dans *amābat* etc.) = o. *-fans* (dans *fufans* « étaient ») = v. irl. *ba*, *bá* « était » (tous issus de **bhw-ā-*) ;

b) à l'exclusion de la syllabation, le prétérit latin *-bat* est identique, au subj. *fuat*, tout comme le v.irl. *ba*, *bá* est identique au subj. *ba* (tous issus de **bh(u)w-ā-* avec une application divergente de la loi de Sievers) ;

c) un prétérit en *-ā-* existe bel et bien en tokharien, mais sa structure diverge de celle de subjonctif en *-ā-* ; *-kälk* « il alla » : *kalkaṣ* « qu'il aille » ;

d) le rapprochement établi par Schindler entre tokh. *käryām* « nous achetâmes » et vieil irl. *-criam* « que nous achetions » se fonde sur l'hypothèse que la forme irlandaise reflète le degré zéro **k^{w}riy-ā-*, là où le subjonctif *-era* « qu'il garantisse » montre que le subjonctif en *-ā-* demandait le degré plein du thème : *-criam* vient donc de **k^{w}rey-ā-* et ne superpose plus au prétérit tokharien ;

e) le rapprochement du lat. *-bat*, de l'o. *-fans* et du v. irl. *ba*, *bá* avec le lit. *bùvo* « il était » (avec *-o* provenant de **-ā-*) est pour le moins spécieux, si l'on considére que ce prétérit fait partie d'un système tout à fait novateur, fondé sur l'opposition entre intransitif en *-o* et transitif en *-ė* (*bùvo* « était » : *nẽšė* « porta ») ;

f) les autres formations prétéritales où *-ā-* apparaît en union stable avec d'autres suffixes, comme c'est le cas pour le lat. *is-ā-* (dans *amāverat*)

ou pour l'arm. **-is-ā-sḱe-* (*gorceac'* «il fit» : **worǵ-is-ā-sḱe-t*), peuvent apporter des éléments de clarification; quant à l'interprétation de Rix (1977), qui sépare le subjonctif en *-ā-* des verbes forts de celui des verbes faibles, et le subjonctif celtique de celui de l'italique – éliminant ainsi le problème du poids que cette formation peut avoir dans le domaine dialectologique –, elle ne va pas plus loin que la pure hypothèse phonétique.

A la lumière des faits examinés plus haut, on peut retenir que cette formation du subjonctif représente une véritable isoglosse qui unit le celtique au latin et aux dialectes italiques; cette isoglosse ne devra cependant pas être interprétée dans un sens généalogique, mais vue plutôt comme le produit d'une contiguïté géographique séculaire.

16

Le morphème du génitif *-ī (pour les thèmes en *-o-*) est attesté dans l'aire celtique soit directement (comme c'est le cas en gallois et en irlandais) soit au moins sous forme de vestiges, comme en corn. *ren verh* «crinière de cheval» (*verh* forme lénifiée de **markī*) et en gall. *Pentyrch*, bret. *Penterc'h* (= irl. *Cenn Tuirch* «Tête de sanglier» : **k^{w}ennos *torkī*). Comme on retrouve le même morphème également en latin (*lupī*), on y a vu un élément de l'unité italo-celtique; cette interprétation ne peut plus être admise aujourd'hui.

Nous ne nous référons pas au fait que ce morphème est absent des dialectes italiques, où **-eís* (morphème de génitif originel des thèmes en *-i-*) a été étendu aux thèmes en *-o-*, par substitution à **-ī* ou à **-osyo*, on ne sait trop. Le point essentiel réside dans le fait que -*ī est attesté aussi en vénète, en messapien et en falisque, se présentant ainsi non pas comme une isoglosse italo-celtique, mais comme une isoglosse spécifique de l'occident.

Au caractère occidental de cette isoglosse semble s'opposer le rapprochement effectué par Wackernagel (1908), entre les formations latines en *-ī* à dépendance verbale (type *parvī facere*) et les formations dites *cvi* de l'ancien indien (type *śuklī karoti/bhavati/syāt* «il rend/devient/soit blanc»). Ce rapprochement, s'il était retenu, ramènerait le morphème en question à l'époque indo-européenne; mais en fait il est très peu convaincant. L'hypothèse la plus probable, en effet, est que les deux types cités plus haut – le type latin et celui de l'ancien indien – n'ont rien de commun du point de vue étymologique (Schindler 1980), puisque les formations *cvi* trouvent leur origine dans l'instrumental des thèmes en *-ī-*, celui-ci étant comparable

non pas au lat. -ī (dans *multī facere*), mais au lat. ē que nous retrouvons dans les formations périphrastiques telle que *rubēfaciō* et *rubēfiō* (avec -ē- issu de **eH*, instrumental des thèmes en *-o-*). L'identité entre lat. *multī facere* et anc. ind. *śuklī karoti/bhavati/syāt* n'est donc qu'apparente.

Il reste encore à éclaircir toutefois la genèse et la fonction originelle du morphème -ī. En effet, sans pécher par généalogisme, il est désormais certain que le latin a possédé autrefois le morphème indo-européen **-osyo* (véd. *-asya*, av. *-ahyā*, arm. *-oy*, gr. -οιο), qui apparaît encore au début de l'époque républicaine dans le *popliosio ualesiosio* (= *Publi Valeri*) du *Lapis Satricanus* et qui était déjà connu au nord de Rome dans le *kaisiosio* (= *Caesi*) de Faleri Veteres (VIIe av. J.-C.). Il est donc clair qu'en latin et en falisque, l'ancien **-osyo* a cédé devant **ī* (Gren-Eklund, 1986), et il est probable qu'il en est de même en celtique ; mais on ne peut pas aller aujourd'hui au-delà de cette constatation.

17

Le rapprochement du futur latin et falisque (mais non pas italique) en *-b-/-f-* (lat. *carebo*, fal. *carefo*) avec le futur irlandais (mais non pas britonnique) en *-f-* (type *léicfid* « laissera ») a été utilisé autrefois comme argument en faveur de l'unité italo-celtique, et aussi comme argument en faveur d'une segmentation ultérieure dans le cadre de cette unité, c'est-à-dire en faveur d'une unité latino-gaélique opposée à une unité symétrique italico-britonnique. En réalité, ce rapprochement se heurte à d'insurmontables difficultés d'ordre phonétique, puisque, alors que le futur latin et falisque contient certainement un élément **-bh(w)-* (comme l'imparfait), ce même élément ne peut être considéré comme la base du futur irlandais en *-f-*, si ce n'est au prix d'un évident tour de force ; le fait même qu'apparaissent périodiquement des travaux destinés à démontrer qu'un *-f-* irlandais peut provenir de **-bh(w)-* (à l'origine en position intervocalique) montre combien cette démonstration est difficile (cf. Quin, 1978, Bammesberger, 1979, etc.).

En réalité, la seule interprétation satisfaisante du point de vue phonétique du futur irlandais en *-f-* est celle de Watkins (1966), qui fait remonter *-f-* à *-sw-ā-* et qui analyse ce futur comme un composé de l'adjectif désidératif en *-su-* et d'un morphème *-ā-* (le même *ā* que dans le subjonctif et que le futur en *-ā-*). Mais cette hypothèse se heurte à deux grosses difficultés : l'adjectif désidératif en *-su-* est propre à l'indien et, qui plus est, présente un redoublement : *ditsu-* « désireux de donner » (**di-dH-su-*). On pourrait

répondre au premier point que celui-ci ne correspond pas à la réalité, si ce type d'adjectif est vraiment à l'origine d'un type de futur en irlandais ; quant à la seconde objection, Watkins soutient qu'en indien aussi il devait exister un adjectif désidératif sans redoublement – identique donc à celui qu'il envisage pour l'irlandais – et dont le dernier et unique représentant serait le véd. *dhakṣu-* (*dah-* « brûler », épithète d'Agni). Mais étant donné que le redoublement apparaît régulièrement dans les formations désidératives, il semblerait plus raisonnable d'interpréter *dhakṣu-* comme une formation en *-u-* sur un thème **dheg^{w}h-s-*, que nous retrouvons aussi en av. *daxša-* « Brand ». Du point de vue étymologique donc, le futur irlandais en *-f-* constitue une énigme non encore éclaircie.

18

Il serait improductif de discuter d'autres formations obscures sur le plan étymologique. Que l'on considère par exemple, le prétérit en *-t* : *-bert* « porta ».

Aujourd'hui, la tendance qui prévaut est celle qui voit dans cette formation un ancien aoriste sigmatique (**bher-s-t*) où, par un processus de resegmentation, le morphème personnel *-t* en est venu à faire partie du thème, donnant ainsi l'origine à des formes comme *-biurt* « je portai » (**bher-st-ō*), *-birt* « portas » (**bher-st-es*) etc. Mais une telle hypothèse a le gros inconvénient de faire considérer comme analogiques beaucoup de ces prétérits : *ét* « protégea », *-sét* « versa » et *-cét* « chanta » ne peuvent dériver *recto itinere* de **emst*, **semst*, **kanst*.

Cela ne veut pas dire qu'il faille revenir à l'ancienne hypothèse d'un aoriste athématique, qui aurait néanmoins quelque fondement dans la comparaison (vieil irl. *-ort* « tua », hitt. *harkt* : **Herg-t-*, v. irl. *-dart* « monta (une vache) », gr. ἔ-ϑορε par métathèse de *ε-ϑεροτ : **(e-)dherH_3-t*). En réalité, la difficulté majeure qui s'oppose à l'une comme à l'autre hypothèse est d'ordre plus général : comment est-il possible qu'on ait conservé en irlandais des groupes consonantiques aussi complexes en position finale ? Est-il phonétiquement admissible qu'un *-bert* puisse représenter un plus ancien **bherst* ou **bhert* ? Tant que l'on ne donnera pas une réponse satisfaisante à ces questions, l'origine du prétérit en *-t* devra demeurer *sub iudice*.

C'est également *sub iudice* que devront rester les hypothèses fondées sur un matériau philologiquement ou exégétiquement incertain ; et la première source d'erreurs est, par définition, le celtique continental. Il suffit

d'un exemple : celui du génitif gaulois *bnanom/mnanom* « des femmes » (Lindeman, 1988).

Le fait que les deux formes apparaissent dans le même texte devrait être déjà inquiétant (que l'on imagine un document grec où l'on aurait aussi bien μνάομαι que βνάομαι ou bien un document en latin où l'on aurait aussi bien *sequere* que *sequese*) ; mais le pire est qu'en réalité aucune de ces formes n'est réellement attestée, puisque dans la magistrale édition du Plomb du Larzac publiée par R. Marichal (1985) nous lisons *bnarcom, semnanom, semnianom* : les deux génitifs gaulois sur lequels travaillent les linguistes ne sont pas des réalités textuelles mais le fruit d'émendations.

19

Bien que le lexique ne signale pas de manière très pertinente les rapports étroits entre les langues (précisément à cause de la facilité avec laquelle des éléments lexicaux peuvent être empruntés par l'une à l'autre), l'examen du lexique des langues celtiques confirme ce que nous avons relevé jusqu'à présent : leur archaïsme général, les affinités particulières avec le latin et les dialectes italiques, le nombre limité d'innovations communes avec d'autres groupes linguistiques.

Des éléments visiblement archaïque ont été déjà mentionnés au § 5 ; en ce qui concerne les rapports avec les langues d'Italie, on relèvera les rapprochements suivants :

v. irl. *scíath* « bouclier », gall. *ysgwyd* : lat. *scūtum*,
gall. *cleddyf* « épée », lat. *gladius*,
v. irl. *tlochur* « parle » : lat. *loquor*,
v. irl. *ad* « loi » : ombr. *arsmor*,
gall. *hoedl* « vie » : lat. *saeclum*,
v. irl. *lestar* « vase » : ombr. *uescla*,
v. irl. *cúl* « le derrière » ; gall. *cil* : lat. *cūlus*,
gall. *ulw* « cendres » : *pulvis*,
gall. *hagr* « sale » : lat. *sacer*.

20

L'existence d'un «lexique occidental» – c'est-à-dire commun au celtique, au germanique, au latin et aux dialectes italiques, avec une éventuelle extension au baltique – est étayée par un petit nombre de lexèmes; les isoglosses que le celtique partage avec le germanique, en tant que témoignages de rapports précis dans le domaine sémantique des activités militaires, sont plus intéressantes; ce sont des isoglosses à caractère exclusif ou, pour le moins préférentielles (ce terme doit être pris dans le sens où, même quand d'autre langues entrent en jeu, il y a tout de même des éléments formels ou sémantiques spécifiques qui lient le celtique au germanique) :

v. irl. *cath* «bataille», gall. *cad* : v. isl. *hǫđ*,
v. irl. *cellach* «querelle» : v. isl. *hirdr*,
v. irl. *bág* «bataille» : v. isl. *bagr*,
v. irl. *trén* «fort» : v. isl. *þrekr*,
v. irl. *neít* «bataille» : got. *ananananþjan*,
V. irl. *níth* «fureur» : v. angl. *níð*,
gaul. *gaison* «javelot» : v. isl. *geirr*,
v. irl. *slacc* «épée» : got. *slahan*,
gall. *llost* «lance» : v. isl. *ljóstr*,
v. irl. *lorg* «massue» : v. isl. *lurkr*,
v. irl. *marc* «cheval» : v. isl. *marr*,
gall. *caseg* «jument» : v.h.a. *hengist*,
v. irl. *fén* «char» : v. isl. *vagn*,
v. irl. *fichid* «combat» : v. isl. *vega*,
v. irl. *calad* «déchaîné» : v.h.a *helid*,
gaul. *dunon* «aire fortifiée» : v. isl. *tún*,
v. irl. *cacht* «prisonnier», gall. *caeth* : v. isl. *haftr*,
v. irl. *giall* «otage» : v. isl. *gisl*,
gall. *budd* «gain» : mba. *būte*,
v.irl. *fotothaid* «effraie» : v. sax. *underbadon*.

Il est évident que ces isoglosses sont dans leur ensemble assez tardives, en ce sens qu'elles présupposent une phase de la technologie militaire considérablement avancée (utilisation de l'épée, de la lance, du char de guerre, etc.); ce sont donc des preuves de rapports accomplis désormais, ou au moins dans une large mesure, dans les sièges historiques des deux groupes.

Il en va certainement de même aussi pour les isoglosses qui concernent les activités artisanales (comme le v. irl. *delg* «épingle», v. isl. *dalkr*) ou

des éléments du paysage géographique (comme le v. irl. *ochtach* «pin», v.h.a. *fiuhta*).

Au regard de cet intense et cohérent faisceau d'isoglosses, les rares termes qui unissent le celtique aux langues de l'aire centrale, comme c'est le cas pour v. irl. *glún* «genou», gall. *glin* : alb. *glu*, n'ont guère d'importance.

Bibliographie
(arrêtée en 1990)

BAMMESBERGER A. (1979). – «On the origin of the Irish *f*-future», *BBCS* 28, 395 *sq.*

BENVENISTE É. (1951). – «Prétérit et optatif en Indo-européen», *BSLP* 47, 11 *sq.*

BERGIN O. (1938). – «On the syntax of the verb in Old Irish», *Ériu* 12, 197 *sq.*

CAMPANILE E. (1966). – «Tre etimologie irlandesi», *SSL* 6, 149 *sq.*

– (1977). – «Sulla preistoria del paradigma di *bé* e *ben* in irlandese antico», *IL* 3, 21 *sq.*

– (1980). – «Un'isoglossa celto-iranica», *SSL* 20, 189.

– (1980a). – «Per l'etimologia di celt. **bardos*», *SSL* 20, 183 *sq.*

– (1991). – «Minima Etymologica», in *Studia Etymologica Indoeuropaea Memoriae A. J. Van Windekens dicata*, ed. cur. L. Isebaert, Louvain, 109 *sq.*

CARRUBA O. (1969). – *Die satzeinleitenden Partikeln in den indogermanischen Sprachen Anatoliens*, Roma.

COWGILL W. (1970). – «Italic and Celtic superlatives and the dialects of Indo-Europeans», in *Indo-European and Indo-Europeans*, ed. by G. Cardona *et al.*, Philadelphia 1970, 113 *sq.*

DILLON M. (1963). – «History of the preverb *to*», *Éigse* 10, 120 *sq.*

EVANS E.D. (1977). – «The contribution of (non-Celtiberian) Continental Celtic to the reconstruction of the Celtic " Grundsprache"», in *Indogermanisch und Keltisch*, hgg. von K.H. Schmidt, Wiesbaden, 66 *sq.*

GREN-EKLUND G. (1986). – «The survival of a hypothesis : Sanskrit *CVI* and Latin gen. sg.-*ī*» in *Festschrift Rundgren*, Stockholm, 157 *sq.*

HAMP E. (1979). – «Indo-European $g^{w}en\text{-}H_{a}$», *KZ* 93, 1, *sq.*

HOFFMANN K. (1967). – *Der Injunktiv im Veda*, Heidelberg.

KELLENS J. (1974). – *Les Noms-Racines de l'Avesta*, Wiesbaden.

KORTLANDT F. (1981). – «More evidence on Italo-Celtic», *Ériu* 32, 1 *sq.*

KRAUSE W., THOMAS W. (1960). – *Tocharisches Elementarbuch*, Band I, Heidelberg.

LANE G. (1959). – «The formation of the Tocharian subjunctive», *Lg.* 35, 157 *sq.*

LINDEMAN F. (1988), «Old Irish *bé* "woman" and some related forms», *EC* 25, 123 *sq.*

MARICHAL R. (1985). – in M. Lejeune, «Le plomb magique du Larzac et les sorcières gauloises», Paris, 10 *sq.* (= *EC* 22, 104 *sq*).

MARSTRANDER C. (1930). – «Tunestenen», *NTS* 4, 294 *sq.*

MEID W. (1972). – «On two points of Celtic morphology», *EC* 13, 346 *sq.*

QUIN E. (1978). – «The origin of the *f*-future : an alternative explanation», *Ériu* 29, 13 *sq.*

RIX H. (1977). – «Das keltische Verbalsystem auf dem Hintergrund des indo-iranisch-griechischen Rekonstrutktionsmodells», in *Indogermanisch und Keltisch* (cité), 132 *sq.*

SCHINDLER J. (1967). – «Tocharische Miszellen», *IE* 72, 239 *sq.*

– (1980). – «Zur Herkunft der altindischen *cvi*-Bildungen» in *Lautgeschichte und Etymologie*, hgg. von Mayrhofer *et al.*, Wiesbaden, 386 *sq.*

SCHMIDT K.H. (1976). – «Zur keltiberischen Inschrift von Botorrita», *BBCS* 26, 335 *sq.*

– (1986). – «The Celtic Languages in their European Context», in *Proc. of the 7th Int. Congress of Celtic Studies*, Oxford, p. 199 *sq.*

STRUNK Kl. (1967). – *Nasalpräsentien und Aoriste*, Heidelberg.

THURNEYSEN R. (1924). – «Keltisches», *IF* 42, 143 *sq.*

VENDRYES J. (1917). – «Les correspondances de vocabulaire entre l'indo-iranien et l'italo-celtique», *MSLP* 20, 265 *sq.*

– (1937). – «Sur le nombre duel», *CRAI* 116 *sq.*

WACKERNAGEL J. (1892). – «Über ein Gesetz der indogermanischen Wortstellung», *IF*, 1, 333 *sq.*

– (1908). – «Genetiv und Adjektiv», in *Mélanges de Saussure*, Paris, 125 *sq.*

– (1930). – *Altindische Grammatik*, Band III (von A. Debrunner), Göttingen.

WAGNER H. (1967). – «Zur unregelmäßigen Wortstellung in der altirischen Alliterationsdichtung», in *Festschrift Pokorny* (hgg. von W. Meid), Innsbruck, 289 *sq.*

– (1982). – «Old-Irish *-bria*, subjunctive of *bronnaid* « injures, damages, spoils, breaks», *ZCP* 39, 83 *sq.*

WATKINS C. (1962). – *Indo-european origins of the Celtic verb*, Dublin.

– (1963). – «Preliminaries to a historical and comparative analysis of the syntax of the Old Irish verb», *Celtica* 6, 1 *sq.*

– (1966). – «The origin of the Celtic *f*-future», *Ériu* 20, 67 *sq.*

– (1976). – «A Hittite-Celtic etymology», *Ériu* 27, 116 *sq.*

CHAPITRE 16

ÉTRUSQUE ET INDO-EUROPÉEN

Dominique BRIQUEL

Si on excepte une soixantaine de gloses, renseignements transmis par des auteurs anciens – qui se réduisent toujours à la traduction d'un mot –, notre information sur l'étrusque provient uniquement des inscriptions qui nous sont parvenues. Celles-ci sont au nombre de 7 000 environ, s'étageant entre le début du VIIe siècle avant J.-C., époque à laquelle l'écriture apparaît en Etrurie, et la fin du Ier siècle avant J.-C., moment auquel le latin supplante définitivement l'étrusque. Ces inscriptions proviennent du pays étrusque proprement dit, c'est-à-dire de la Toscane, et également d'autres régions où les Etrusques se sont établis à certaines périodes, comme la plaine padane vers le nord, et vers le sud, la Campanie et même Rome ; certaines trouvailles plus lointaines, dues à l'importante activité maritime des Etrusques, sont à signaler : Carthage, la Corse et le littoral du Languedoc ont livré des inscriptions.

Ce chiffre peut sembler relativement important : à titre de comparaison les inscriptions latines de la période correspondante (c'est-à-dire époques royale et républicaine) ne sont que de l'ordre de 3 000. Mais l'énorme majorité de ces inscriptions n'offre que des énoncés très courts : on ne dispose que de trois textes un peu longs, respectivement de 1 200,

300 mots (dans ces deux cas il s'agit de calendriers religieux, indiquant des cérémonies à effectuer certains jours ; l'un est écrit sur lin et a été trouvé bizarrement en Egypte, où il avait servi à envelopper une momie ; il est appelé « livre de Zagreb » d'après son lieu de conservation ; l'autre est une inscription sur une sorte de tuile, provenant de Capoue, d'où son nom de « tuile de Capoue »), et 130 mots (un jugement d'arbitrage dans une affaire de délimitation de propriétés, porté sur une borne, le « cippe de Pérouse »). Le sol de l'Etrurie n'a pas livré l'équivalent de ce que représentent pour l'ombrien les tables de Gubbio, ces sept plaques de bronze portant de longues prescriptions rituelles, ou même pour l'osque la loi municipale trouvée à Bantia, texte dont sont conservés près de 400 mots. La plupart de ces inscriptions étrusques est de caractère funéraire – et fournit donc des indications limitées (nom du défunt, avec éventuellement ses liens de parenté, son âge, quelquefois des indications sur sa carrière). D'autres catégories d'inscriptions bien représentées ont un aspect encore plus stéréotypé : nous disposons de nombreuses marques de possession, avec un énoncé du type « j'appartiens à un tel », exprimé en étrusque par le pronom de première personne au cas sujet *mi* suivi du nom du possesseur au génitif, ou de dédicaces, avec une formule du genre de « un tel m'a donné à un tel », où l'on a le pronom de première personne au cas objet *mini*, un verbe signifiant « a donné » (par exemple *muluvanice* ou *turice*), le nom du dédicant au cas direct et le nom du dédicataire à un cas oblique qu'on peut appeler datif. On voit d'emblée que cette langue ne nous est accessible que par une documentation réduite : toute la littérature étrusque dont nous parlent les auteurs grecs et latins, et qui semble avoir été particulièrement développée dans le domaine religieux, est perdue, et nous ne pouvons nous en faire une certaine idée que par des documents exceptionnels comme les calendriers sacrés auxquels nous avons fait allusion.

Mais le problème que pose l'étrusque n'est bien sûr pas seulement celui que pose l'extension limitée de notre documentation. L'étrusque reste une langue qui nous échappe et la plupart des documents, surtout ceux de plus grande longueur et d'énoncé moins stéréotypé, nous sont incompréhensibles. Selon la remarque que faisait déjà à l'époque d'Auguste un historien grec, Denys d'Halicarnasse, cette langue est isolée et ne ressemble à aucune autre. Malgré des tentatives innombrables, faites dans toutes les directions possibles, on n'est parvenu à expliquer l'étrusque par aucune langue-sœur, et on ne peut pas ranger cette langue dans une famille de parlers connue. En fait les seuls idiomes qui paraissent lui être clairement apparentés sont eux-mêmes des parlers mal attestés et qui échappent à une véritable analyse. Il s'agit, d'une part, du lemnien, qui était parlé au VIe siècle avant J.-C. dans l'île de Lemnos, dans le nord de l'Egée – connu par une épitaphe d'une trentaine de mots –, et d'autre part, du rétique, attesté par une petite centaine de courtes inscriptions du pays des Rhètes, dans les Alpes du Nord de l'Italie.

Il est évident, dans ces conditions, que tout essai de description de l'étrusque se heurte à ces difficultés majeures que sont et la limitation de notre *corpus*, et l'obscurité persistante de la langue. La première fait que certaines formes de la langue ont *a priori* peu de chances d'être rencontrées dans le type de documentation que nous avons : par exemple on peut douter que s'y trouvent des verbes à la deuxième personne du singulier, ou à une forme exprimant le futur. Des pans entiers de la morphologie nous restent de ce fait hors d'atteinte. Quant à l'obscurité de la langue, les conséquences en sont encore plus obvies : il existe fort peu de points qu'on peut estimer vraiment assurés, et toute description de la langue est donc inévitablement amenée à faire une part à l'hypothèse. Celle-ci peut être bien sûr plus ou moins grande, mais c'est un état de choses dont il est nécessaire d'être averti dès qu'on aborde tout essai de présentation de la langue !

1

Au moins les inscriptions étrusques ne posent pas de problème de lecture : les Etrusques usaient de l'alphabet grec de type occidental que leur avaient transmis les colons helléniques établis à Cumes en Campanie au VIII[e] siècle avant J.-C., et cet alphabet étrusque, emprunté par les Latins, est à l'origine de celui que nous utilisons. De ce fait nous pouvons nous faire une idée du système phonologique de la langue. Or sur ce point déjà l'étrusque offre des caractères remarquables, qui paraissent se retrouver en lemnien et en rétique :

a) la langue possède une série d'occlusives sourdes (labiale [p], dentale [t], vélaire [k]) et une série d'occlusives aspirées ([ph], [th], [kh]), mais n'offre pas la série sonore homologue ([b], [d], [g]) : les lettres grecques correspondantes, conservées dans l'alphabet théorique, ne se rencontrent jamais dans les inscriptions – sauf l'ancien *gamma* grec, qui est devenu signe de sourde et non plus de sonore, donnant ainsi naissance à notre C ;

b) elle n'offre qu'une voyelle d'arrière, notée par la lettre U (mais par O en lemnien) – alors qu'elle possède par ailleurs un [a], un [e], et un [i] ;

c) il existe deux sifflantes distinctes (notées respectivement par le *sigma* et la lettre dite *san* – en forme de M – de l'alphabet modèle grec, avec d'ailleurs une inversion des valeurs entre le Nord et le Sud de l'Etrurie) ; on discute sur la valeur exacte de cette opposition (opposition de forte à douce ou de sifflante normale à chuintante). Par ailleurs on rencontre la lettre Z, qui doit noter une affriquée [ts].

On peut relever en outre que l'étrusque, comme les parlers italiques, possède un phonème [f]. Le fait qu'il ait été rendu primitivement dans l'écriture par le digramme *digamma* + H (FH), avant qu'on ne crée pour lui un signe spécial, en forme de 8, et aussi celui qu'il alterne fréquemment avec [h] suggère qu'il a un caractère bilabial.

Ce système phonologique semble être resté stable tout au long de l'usage écrit de l'étrusque : on ne note pas en étrusque, comme par exemple en ombrien, l'apparition de signes nouveaux ou la transformation de la valeur de signes anciens pour rendre compte de l'évolution de la langue. Mais cela ne veut pas dire que la langue n'ait connu pendant ces sept siècles aucune évolution d'ordre phonétique. Si un phénomène comme le rhotacisme, qui touche l'ombrien, le falisque autant que le latin, n'apparaît pas en étrusque, les effets d'un accent d'intensité initiale y sont fort sensibles : on constate vers les débuts du V[e] siècle avant J.-C., après l'altération du timbre des voyelles en position intérieure, leur disparition – ce qui permet de distinguer un état « récent » de la langue, à partir du IV[e], par opposition à un état « archaïque », celui des plus anciennes inscriptions, et un état « classique », intermédiaire. Ainsi le verbe « a donné » qui s'écrivait *muluvanice* au VII[e] siècle se rencontre au siècle suivant sous les formes *mulvenece* ou *mulvunuce*; l'autre verbe de même sens, attesté sous les formes alternatives *turice* ou *turuce* au VI[e] siècle, est ensuite réduit à *turce*. Cela se traduit par des syncopes et des disparitions de phonèmes parfois fort importantes : ainsi le prénom féminin qui se lit *ranθu* (mais aussi *ravnθu*) dans les inscriptions récentes représente la réduction de l'ancien *racvenθu*. Mais plus généralement on aboutit à des mots présentant, en position interne, des accumulations de consonnes qui sont un aspect caractéristique des formes de l'étrusque plus récent : ainsi le grec Ἀλέξανδρος se voit rendu par *elχsntre*.

2

Le nom étrusque ne semble pas connaître de distinction de genre. Une opposition masculin/féminin n'apparaît clairement que dans le cas particulier de l'onomastique. Dans ce cas la forme du gentilice indique s'il s'agit d'un homme ou d'une femme; cette dernière a un gentilice qui adjoint un suffixe *-i* à la forme masculine, par exemple *pumpui* en face de *pumpu*; de ce fait, beaucoup de gentilices étant formés par une suffixation en *-na* leurs formes féminines seront en *-nai* (par exemple fém. *tarχnai* à côté de *tarχna*, correspondant au latin Tarquinius) qui évolue

vers *-nei* et *-ni* (par exemple *spurinei* et *spurini* à côté de Spurina, nom d'une importante famille de Tarquinia qui devait fournir un haruspice à César). Mais c'est là un phénomène strictement limité : même dans le cas des prénoms, le procédé n'est employé que pour déterminer des formes féminines là où elles sont en parallèle avec des formes masculines (par exemple fém. *larθi* en face du masc. *larθ*, ou avec un autre suffixe fém. *velelia* en face du masculin *venel*) ; dans des prénoms féminins courants comme *θana* ou *θanaχvil*, rien ne montre, formellement, qu'il s'agit de désignations de femmes. Et en dehors de l'onomastique, mis à part de rares couples (comme *lautniθa*, affranchie, en face de *lautni*, affranchi), rien ne vient caractériser un nom comme ayant un caractère féminin plutôt que masculin : ainsi le nom de la fille, *seχ* n'offre rien qui le distingue comme étant d'un genre différent de celui du fils, *clan*.

Certains traits de la flexion nominale étrusque se laissent déterminer : ce qui ne veut pas dire que nous soyons en mesure de dresser un tableau complet du paradigme, ni même une simple liste des cas existants dans la langue ! Le cas-sujet ne se distingue pas du cas-objet pour les noms – ce qui est différent du cas des pronoms (nous avons déjà évoqué l'opposition *mi/mini* pour le pronom personnel de première personne). Le nom de la tombe est en étrusque *śuθi* : on le trouve aussi bien dans un formulaire où il accompagne une forme de cas-sujet de pronom démonstratif *eca* (par exemple *eca śuθi anes culcnies*, c'est dire « ceci (est) la tombe de Ane Culcnie ») que le cas-objet *cn (vel matunas larisalisa an cn śuθi ceriχunce*, « Vel Matunas, celui de (c'est-à-dire fils de) Laris, celui-ci (avec une sorte de relatif de liaison *an*) a fait cette tombe »).

L'existence d'un génitif est bien assurée. Nous l'avons déjà évoqué à propos des formules de possession avec *-mi*. Nous pouvons prendre pour exemple *mi tites latines*, « je (suis) à » – ou plus exactement « de » – Titus Latinus : la transcription étrusque de ce nom latin serait au cas-direct (cas-sujet ou objet) *tite latine*. Il est aussi très fréquent dans les épitaphes, en fonction de complément de nom de termes comme *clan* (« fils »), *seχ* (« fille ») ou *puia* (« épouse ») : ainsi une Larθi Spantui sera désignée comme fille de Larce Spantu (*larces spantus seχ*) et épouse de Arnθ Partunus (*arnθal partunus puia*). Mais dans de telles désignations le simple génitif peut être remplacé par une forme en *-śa*, qui est à considérer comme formée à l'aide d'une sorte de pronom enclitique suffixé (et qui peut à son tour se fléchir et être mise au génitif-*śla*) ; nous avons vu que Vel Matunas était dit *larisalisa* : on aurait pu tout aussi bien avoir *larisal clan*.

Morphologiquement ces génitifs sont en *-al* ou en *-s*. Sans que le principe de répartition soit toujours très clair, on peut dire que la forme en *-al* est de règle pour les noms féminins se terminant en *-i* (alors que des prénoms féminins ayant d'autres terminaisons ont leur génitif en *-s* : *ramθa* fait *ramθas*, *θanχvil θanχvilus*) et pour des noms, y compris des éléments onomastiques masculins, se terminant au cas direct en *-s* ou en *-θ*.

Comme illustration nous pouvons citer l'inscription *larθ arnθal plecus clan ramθasc apatrual*, où ce Larθ est dit fils d'Arnθ Plecu et (copule *-c*) de Ramθa Apatrui. Lorsque le génitif en *-s* apparaît pour des termes ayant au cas direct une consonne finale, on trouve généralement une voyelle devant la sifflante. Ce n'est pas toujours le cas (le nom du fils *clan* fait *clens*, avec un changement de timbre vocalique difficile à expliquer) et cette voyelle est alors de timbre variable : on a parfois supposé que c'était la trace d'une voyelle terminant le radical qui aurait disparu au cas direct. Il faut en outre compter avec des formes de génitif différentes, attestées dans les plus anciennes inscriptions : à la place de formes en *-al*, on y rencontre des formes en *-ia* ou *-aia*. Ainsi sur une fibule en or de Chiusi du VII[e] siècle avant J.-C. on lit *mi araθia velaveśnaś*, et de même un vase de Vulci de cette époque porte la marque de possession d'une Ram(u)θa Kansinai sous la forme *mi ramuθas kansinaia*, et non *mi ram(u)θas kansinal* comme on le trouverait par la suite.

Un autre cas qui se laisse assez bien repérer est un cas en *-si* ou *-ale* (plus anciennement *-ala*), qui peut se définir formellement comme obtenu par l'adjonction d'une voyelle au cas précédent. On lui donne généralement le nom de datif : il joue un tel rôle de datif dans les formules de don, avec *mini muluvanice* ou *turice*, que nous avons évoquées. Ainsi *mini spuriaza* [ ]*rnas muluvanice alśaianasi* signifiera « Spuriaza [ ]rnas m'a donné à Alsaiana », avec une forme de datif *alśaianasi* différente du génitif qui serait *alśaianas ; mini muluvanice mamarce apuniie venala* sera « Mamercus Aponius m'a donné à Venai » (forme de prénom ou nom individuel féminin), avec un datif en *-ala* alors que le génitif serait *venal*. Mais, si on emploie ce terme de datif – ce qui nous semble toujours admissible –, il faut avoir conscience de ce que la zone d'emploi de ce cas ne recouvre pas celle du datif dans une langue comme le latin. Ce cas en *-si* ou en *-ala/ale* apparaît dans des tournures avec des formes verbales au passif (que nous aurons à examiner), où il doit avoir une fonction de complément d'agent. Dans la « signature » que porte un vase au riche décor gravé *mi araθiale ziχuχe*, à côté du verbe passif *ziχuχe*, formé sur une racine signifiant écrire, et ici graver, *araθiale*, « datif » du prénom Ar(a)nθ (qui serait théoriquement plutôt *aranθale* : mais il faut compter avec divers phénomènes comme une palatalisation de la dentale), désigne certainement le graveur de ce vase, et il faut comprendre « j'ai été gravé par Ar(a)nθ ».

Mais la question est compliquée par le fait qu'on rencontre également dans de tels contextes, avec des expressions verbales de type passif, et au moins à époque récente des formes qui ne sont pas purement et simplement ces cas en *-si* ou *-ala/ale* (qui eux sont attestés surtout, mais pas exclusivement, dans des inscriptions anciennes). Nous pouvons citer ainsi une inscription tombale de Tarquinia où on lit *larθ tutes anc farθnaχe veluis tuteis θanχviluisc turialsc*, ce qui peut se traduire par « Larθ Tutes ; et (*-c*) celui-ci fut engendré (*farθnaχe*) par Vel Tutes et Θanaχvil Turiei ».

Le verbe passif est ici accompagné d'un cas oblique qui n'est ni le génitif (qui serait *velus tutes* et *θanχvilus turial*), ni ce datif que nous avons examiné (qui serait *velusi tutesi* et *θanχvilusi turiale*). Faut-il alors admettre l'existence d'un autre cas, autonome, qui serait un ablatif ? Il serait caractérisé, pour les thèmes à génitif en *-s*, par une sorte de palatalisation de la voyelle précédente (*ui* au lieu de *u*, *ei* au lieu de *e*, on aurait ainsi *e* au lieu de *a* dans *ramθes*, attesté dans le même contexte par une autre inscription), et pour ceux à génitif en *-al*, par une terminaison *-als* connue aussi dans d'autres contextes (ainsi *truials* comme didascalie d'une peinture, signifiant le Troyen, celui qui est (originaire) de Troie). Mais à cette explication on en a parfois préféré une autre : que ce cas ait résulté d'une évolution récente de l'ancien cas en *-si* ou *-ala/ale*, par évolution phonétique (*-si* altéré en *-s* palatalisant la voyelle précédente) et réfection morphologique (adjonction de *-s* à un *-al* ne se distinguant plus du génitif). La question, on le voit, est loin d'être claire.

Elle l'est encore moins pour d'autres cas. L'existence de formes de cas obliques en *-e*, fréquents dans un texte comme celui de la momie de Zagreb, paraît assurée : mais ces formes ne se laissent guère analyser. On rencontre aussi des formes de locatif (qui peuvent aussi avoir valeur temporelle), en *-θi* (par exemple dans l'expression *unialθi*, qui est formée sur le génitif du nom de la déesse Junon étrusque, Uni, et signifie « dans le temple de Junon »), mais aussi en *-θ* (par exemple dans *śuθiθ*, dans la tombe), en *-ti* ou encore simplement en *-i*, ainsi dans *zilci*, qui coexiste avec la forme *zilcti* – ces mots apparaissant dans des formules de datations par des magistrats éponymes (*zilc* étant une désignation de magistrat), comme *zilci velusi hulcniesi*, sous la magistrature de Vel Hulcnie. Le cas en *-e* lui-même a parfois une fonction de locatif : ainsi *capue* signifie à Capoue dans un texte que nous aurons à analyser. Là encore, si cette fois la fonction est claire, la distribution de ces formes ne l'est guère.

Nous avons laissé de côté la catégorie du nombre. Or elle se laisse quelque peu appréhender en étrusque. On peut reconnaître au moins deux manières de former des pluriels, qui ne paraissent pas utilisées pour les mêmes mots, dans la mesure du moins où notre documentation permet de trancher. L'une consiste à utiliser un suffixe, de sens peut-être collectif, *-χva*. Citons ainsi une formule prescriptive du livre de Zagreb : *celi huθiś zaθrumis flerχva neθunsl śucri θezeric*, ce qui signifie à peu près « le vingt quatrième (si *huθ*, qui est un chiffre, vaut quatre, ce qui est discuté) jour du mois de Celi, il faut offrir et sacrifier (avec des sortes de « gérondifs » en *-ri*) au dieu Neptune des offrandes *fler* » ; *flerχva* forme ici un pluriel du *fler* ailleurs attesté. Mais on rencontre également des pluriels formés par l'adjonction d'un suffixe en *-r* : *aiser* signifie « les dieux » à côté du singulier *ais, clenar* « les fils » à côté du singulier *clan*. Et ce suffixe se retrouve à tous les cas du pluriel, le morphème du cas, le même qu'au singulier, venant alors s'adjoindre à lui. On peut suivre à peu près le

fonctionnement du système pour le nom du fils, *clan*. Au singulier le génitif est *clens* (adjonction du morphème *-s*), le datif *clensi* (morphème *-si*), le pluriel sera formé par l'adjonction, avant la terminaison casuelle, de *-ar* : on aura au cas direct *clenar*, au génitif *cliniiaras* (avec des palatalisations peu claires et l'adjonction d'une voyelle devant le *-s*) et au natif *clenarasi*. On a un procédé qui fait penser au système des langues agglutinantes.

Le verbe est surtout connu par des formes de troisième personne du parfait – ce qui est normal dans une documentation épigraphique qui offre beaucoup d'exemples d'expressions signifiant en gros « un tel a donné » (*muluvanice, turice, alice*), « a fait » (*zinace, menece, ceriχunce*) ou « a écrit » (*ziχunce*). De ce fait une terminaison de parfait en *-ce* (écrit aussi *-ke*) est bien attestée. Mais elle s'emploie autant pour le pluriel que pour le singulier : dans l'inscription tombale *laris avle larisal clenar sval cn śuθi ceriχunce*, on a deux sujets, les frères Laris et Aule, fils de Laris, qui « de leur vivant (*sval*) ont fait cette tombe ».

Un acquis récent concerne l'existence d'une opposition entre un parfait actif, avec suffixe *-ce*, et un parfait passif, avec suffixe *-χe*. A la signature d'artiste active *metru menece*, « Metron (nom grec Μέτρων) a fait » s'oppose la tournure passive *mi titasi cver menaχe*, « j'ai été fait comme cadeau (*cver*) par Tita (avec datif en *-si*) ». Dans ce cas on a affaire à une première personne. Mais nous avons des exemples où cette désinence en – *χe* correspond à des troisièmes personnes. Il n'est pas impossible que le verbe étrusque ne connaisse pas de distinction entre ce qui relèverait pour nous de la première et de la troisième personne (et nous avons vu qu'il ne distinguait pas entre singulier et pluriel pour la troisième personne). Il est vrai qu'on a voulu au contraire reconnaître une désinence spécifique de première personne dans des formes en *-un* (et, notons le, soit en *-cun*, soit en *-χun*). On a en effet proposé de voir une forme de première personne sur une lamelle de plomb portant une formule d'exécration, dans le terme *θapicun* qui y revient plusieurs fois et serait un verbe signifiant « je maudis », ou peut-être « j'ai maudit »; et un passage du texte de la momie de Zagreb, malheureusement peu clair, présente un *slapiχun* : le morphème de première personne serait alors en *-un*. La question reste, on le voit obscure.

Pour ce qui est du passif, à côté de formes avec le verbe en *-χe* on rencontre fréquemment des formules de sens analogue où le verbe n'est pas exprimé à un mode personnel, mais par une sorte de nom verbal en *-u* : en alternance avec les formules de don du type *mini muluvanice*, on en rencontre du type *mi mulu* avec cas en *-si* ou *ala/ale*, à fonction d'agent (par exemple *mi mulu kaviiesi*, « j'ai été donné par Gavius »). Et cette forme en *-u* peut revêtir tout autant à nos yeux un sens actif qu'un sens passif : à côté d'exemples de sens passif comme *mulu, turu* ou *aliqu* (au sens de « donné »), on rencontre, pour un verbe intransitif, *zilaχnu* (ayant occupé la fonction de magistrat *zilaχ*), ou *zicu* (employé comme

cognomen, équivalant au latin Scriptor, signifiant «l'Ecrivain») pour le verbe actif *ziχ*, écrire – avec une perte de l'aspiration due au caractère tardif de l'inscription.

Les autres formes de la flexion verbale sont moins aisément déterminables. Il semble qu'au morphème en *-ce* de parfait s'oppose un morphème parallèle en *-e* pour le présent : deux dédicaces jumelles d'Ischia di Castro *mine muluvenice avile acvilnas* et *mine muluvene avile acvilnas* seraient donc à comprendre comme exprimant le don l'une au parfait et l'autre au présent («Aulus Acvilnas – équivalent du latin Aquilius – m'a donné / me donne»), et de même *ame* serait le présent «il est» en face du parfait *amce* «il a été». On a même proposé d'étendre cette distinction aux noms verbaux en *-u* : ceux en *-u* auraient une valeur de présent, et ceux en *-cu* (ou *ku* ou *qu*) de parfait : effectivement on a un exemple – malheureusement peu clair – de *alu* à côté du fréquent *aliqu* (ou *alcu*). Autre document de type exceptionnel, le cippe de Pérouse porte des forme verbales en *-a*, comme *ama* que l'on est porté à rapprocher de *ame/amce*, il est, il fut : dans ce texte juridique, il semble s'agir de subjonctifs à valeur de prescription. Mais dans la momie de Zagreb, nous l'avons vu, les prescriptions sont plutôt exprimées par ce qu'on a appelé un «gérondif», en *-ri*. Et l'on trouve aussi un impératif, réduit au simple thème verbal : *tur*, dans les prescriptions qu'offre ce texte, où il alterne avec le «subjonctif» *tura-*, doit signifier «donne».

On repère un peu mieux des exemples de participes. On rencontre des formes en *-as*, ou, équivalentes en *-asa* (par exemple sur des dédicaces *aleθnas v. clenar... ci acnanasa*, «Vel Aleθnas ayant eu trois fils», et *metli arnθi... ci clenar acnanas*, «Arnθi Metli – une femme – ayant eu trois fils»). Parfois le suffixe est plus complexe : *-θas* (ou *-θasa*). On en a deux exemples, pour les verbes *ten-*, signifiant occuper une fonction, et *sval-*, signifiant vivre, dans une épitaphe de Tarquinia : *velθur partunus larisaliśa clan ramθas culcnial zilχ ceχaneri tenθas avil svalθas LXXXII*, «Velθur Partunus, fils de Laris, fils de Ramθa Culcniei, ayant occupé les fonctions de *zilχ ceχaneri*, ayant vécu 82 ans».

Nous avons déjà rencontré des exemples de pronoms, comme le pronom personnel de première personne *mi/mini*, le relatif *an*, qui fonctionne aussi comme relatif de liaison, ou des démonstratifs. Ceux-ci sont assez nombreux, et on peut considérer qu'il existe une série en *-ta* et une série en *-ca*, avec éventuellement une voyelle préposée (*eca/ica* ou *ita*). Il est probable qu'existe entre elles une distinction entre objet rapproché et objet éloigné, comme celle qui serait alors posée entre les deux objets désignés dans la dédicace sur une lamelle d'or découverte dans le temple de Pyrgi *ita tmia icac heramasva vatieχe*, «cet espace sacré (?)-là et ces statues (?)-ci (avec pluriel en *-cva* réduit à *-va*) ont été consacrées (?)». Ces pronoms peuvent de plus se combiner enclitiquement avec des noms, ou des formes nominales : on en a un bon exemple dans l'épitaphe d'un Etrus-

que qui avait dû combattre lors de la seconde guerre punique, avant de mourir à l'âge respectable de 106 ans : *felsnas la leθes svalce avil CVI murce capue tleχe hanipaluscle*, «Larθ Felsnas (fils de) Leθe vécut 106 ans, fit une action (verbe actif *murce*) à Capoue, subit une action (verbe passif *tleχe*) de la part des gens d'Hannibal»; *hanipaluscle* est formé sur le génitif du nom d'Hannibal, auquel est adjoint le démonstratif *-ca*, ici à un cas oblique- *cle*. La déclinaison de ces pronoms est plus complexe que celle des noms, et distingue du cas sujet un cas objet en *-n*, qui semble être la réduction d'un ancien *-ni*. On trouve aussi des génitifs aussi bien en *-s* (type *cs*) qu'avec *-l* (type *cla*), des cas obliques en *-e* (type *cle*) que des locatifs (type *clθi*). Mais l'exemple de Pyrgi nous montre que la distinction singulier/pluriel ne se laisse pas repérer.

Parmi les mots invariables, nous avons déjà rencontré la copule enclitique *-c*; on trouve également un enclitique *-m*, dont le sens, peut-être légèrement adversatif, n'est en fait guère différent : par exemple *ramθa matulnai seχ marces matulnas puiam amce seθres ceisinies*, «Ramθa Matulnai, fille de Marce (Marcus) Matulna, et (*-m*) elle fut épouse (*puia*) de Seθre Ceisinie»; il peut se combiner avec une sorte de forme pronominale figée, pour donner la conjonction de coordination *etnam*. On peut également citer un *iχ*, qui signifie «comme». Ce sens est assuré pour une formule comme celle qui termine le texte du cippe de Pérouse *iχ ca ceχa ziχuχe*, «comme cela a été écrit au-dessus (*ceχa*)» et qui a son exact équivalent dans des conclusions de textes juridiques latins ou ombriens. On a récemment proposé de reconnaître dans un *ei* une particule à sens négatif, mais qui serait plutôt le correspondant du *ne* latin que du *non*, et donc serait une marque de défense et non pas la simple négation. On peut également signaler certaines prépositions, ou postpositions : c'est le cas de ce *ceχa*, qui a certainement la valeur de «pour» dans la formule *clen ceχa*, qui apparaît dans des dédicaces d'objets faites par des femmes pour demander à un dieu de leur accorder un fils (ou pour le remercier de le leur en avoir accordé). Mais dans bien des cas le sens est loin d'être clair : c'est ce qui arrive pour un *-pi* et un *-ri* enclitiques.

3

Etant donné l'obscurité de la langue, c'est un truisme que de dire que le vocabulaire étrusque ne nous est que très imparfaitement connu. Et c'est sans doute là que nous nous heurtons le plus au problème de l'isolement de la langue. On peut en prendre comme illustration la série

des chiffres de un à six – connue par des dés à jouer marqués du nom des chiffres. Elle ne se laisse pas ramener à des parallèles attestés : on a *θu, zal, ci, huθ* (mais d'autres estiment que « quatre » est *śa*), *maχ, śa* (à moins que « six » ne soit *huθ*). Depuis longtemps on a rapproché le nom de nombre *huθ* du toponyme attique Ὑττηνία, qui désignait une Tétrapole, confédération de quatre bourgades. Et cette analogie rentre dans une série de faits parallèles (nom de la femme en étrusque *puia*, et grec ὀπυίω « épouser », nom du héros étrusque Tarchon et grec ταρχύω « rendre les honneurs funéraires », et théonyme hittite *Tarhundas*). Ces termes, en grec, appartiennent au substrat préhellénique : il semble par là que l'étrusque offre des traits qui permettent de le relier au substrat linguistique pré-indo-européen, méditerranéen, tel qu'on le décèle par exemple par ces traits préhelléniques du grec. Mais les faits sont complexes. L'étrusque a été au moins fortement influencé par l'indo-européen. Ainsi, dans la série des noms de parenté, nous avons vu le nom de la femme, *puia*, qui peut être qualifié de « méditerranéen », les noms du fils et de la fille, *clan* et *seχ*, qui n'appellent pas de rapprochements avec des langues connues ; d'autres noms, celui du père *apa*, celui de la mère *ati* (qui peut être considéré comme un féminin en *-i* par rapport à un *ata*), celui du petit-fils *papals*, sont formés sur des thèmes comme *apa/papa/ata*, qui se rencontrent dans divers groupes linguistiques et ne peuvent de ce fait pas avoir de signification pour déterminer des apparentements ; mais on rencontre aussi *nefts*, qui est clairement un homologue du latin *nepos*, génitif *nepotis*. Et, plus significativement que ce trait qui peut tenir à un emprunt, nous avons rencontré des points du système de la langue qui peuvent, avec plus ou moins de probabilité selon les cas, tenir à des similitudes structurelles entre l'étrusque et les langues indo-européennes : citons par exemple le parfait en *-ce*, la copule enclitique *-c*, le pronom personnel *mi*, la désinence de locatif *-θi*. A ces analogies, il est vrai, on peut opposer des traits qui semblent foncièrement différents – comme les éléments de type agglutinant que nous avons relevés dans la flexion nominale. C'est pourquoi on a forgé, pour l'étrusque, le concept de « péri-indo-européen » ; cela signifierait que l'étrusque serait une langue non indo-européenne, correspondant à ce qui était parlé dans le bassin méditerranéen, ou au moins une de ses parties, avant l'établissement des parlers indo-européens dans cette zone, mais qui se serait développée à proximité de tels parlers et en contact avec eux : cette situation expliquerait certaines homologies dans la morphologie qui peuvent difficilement passer pour des faits d'emprunt.

Bibliographie

BONFANTE G. et L. (1983). – *The Etruscan Language, an Introduction*, Manchester.

CRISTOFANI M. (1973). – *Introduzione allo studio dell'etrusco*, Florence.

PALLOTTINO M. (1978). – *La langue étrusque*, Paris.

PFIFFIG A.J. (1969). – *Die etruskische Sprache*, Graz.

RIX H. (1984). – Chapitre « La scrittura e la lingua », dans M. Cristofani éd., *Gli Etruschi, una nuova immagine*, Florence, p. 210-238.

STACCIOLI R.A. (1977). – *Il « mistero » della lingua etrusca*, Rome.

LOUIS-JEAN
avenue d'Embrun, 05003 GAP cedex
Tél. : 92.53.17.00
Dépôt légal : 559 — Juillet 1994
Imprimé en France

La collection **"Sciences du Langage"** se propose de publier des travaux analytiques et des monographies, aussi bien que des synthèses. Elle est concernée par tous les domaines des sciences du langage et est ouverte à toutes les écoles. Elle privilégie la lisibilité, l'innovation, l'achèvement.

Les Dialectes indo-européens d'A. Meillet remontent à 1908. Il nous a semblé utile qu'un ouvrage français dressât vers la fin de ce siècle un état de la question sur ces dialectes, dont certains n'étaient pas encore connus de Meillet, tandis que la problématique de ceux qu'il traite a été renouvelée. Notre titre, *Langues indo-européennes*, implique qu'il y a des absents : le latin et le grec, que nous supposions mieux connus du public cultivé non spécialiste que les autres langues : tokharien, indo-iranien, langues anatoliennes comme le hittite, arménien, thrace, albanais, balte et slave, germanique, italique, celtique.
Le livre, réunissant d'éminents spécialistes internationaux, aborde également sur la reconstruction culturelle et linguistique, la méthode comparative, les noms des Aryens et est complété par un chapitre sur l'étrusque, non apparenté jusqu'à plus ample informé aux langues indo-européennes, mais dont la comparaison typologique avec ces dernières pose d'intéressants problèmes.

Françoise Bader, qui a coordonné l'ouvrage, est directeur d'études de grammaire comparée des langues indo-européennes à l'Ecole Pratique des Hautes Etudes à Paris.



PRIX : FF 350
ISSN 0991-5877
ISBN 2-271-05043-X